CAMBRIDGE

A-LEVELS

FURTHER

MATHEMATICS

9231

NOTES

WRITTEN BY: DEVIN KY

1. Table of Contents

Paper 1 Topics

Chapter	Topic	Page
1	Polynomials and Rational Functions	5
2	Polar Coordinates	20
3	Summation of Series	29
4	Mathematical Induction	36
5	Differentiation and Integration	43
6	Differential Equations	66
7	Complex Numbers	77
8	Vectors	85
9	Matrices and Linear Spaces	97

Paper 2 Topics

Chapter	Topic	Page
10	Momentum and Impulse	129
11	Circular Motion	138
12	Equilibrium of a Rigid Body under Coplanar Forces	146
13	Simple Harmonic Motion	160
14	Rotation of a Rigid Body	169
15	Further Work on Distributions	180
16	Inference using Normal and t-Distributions	193
17	Chi-squared Tests	207
18	Bivariate Data	216

2. Pre-Requisite Knowledge

Paper 1: Pure Mathematics

This part of the notes is written with the assumption that the reader completed the contents of Mathematics Paper 1 (P1) and Paper 3 (P3). The reader can start with any topic after finishing the module P3, but it is advised to start with Topics 3, 5, 8 or 9 they introduce new notations that might be used in other topics.

Paper 2: Mechanics and Statistics

This part of the notes is written with the assumption that the reader completed the contents of Mathematics Paper 4 (M1) and Paper 6 (S1). Contents of Mathematics Paper 5 (M2) and Paper 7 (S2) that are required are discussed with several examples. The reader can start with Chapters 10 to 14 after M1 and Chapters 15 to 18 after S1. However, it is recommended that the reader also finished P3 as questions may require concepts like exponents and calculus techniques. Readers are advised to read the Mechanics or Statistics parts in ascending order as the pre-requisite knowledge is gradually covered when required for progress.

3. Answering and studying techniques: some Do's and Don'ts

DO'S

1. Read and familiarise yourself with the syllabus and the format of examination. I understand that some students take up FM by themselves so it's important to know what to do to prepare for it. Make sure you know what concepts are involved and how much work you have to do in the time limit.
2. Manage your time properly. Make sure you fully utilise the three hours given to answer all the questions and only attempt both alternatives if you have a lot of time left. Sometimes questions that require a lot of thinking and reasoning (especially Mechanics) might take a lot of time. If you find yourself spending more than 30 minutes on a question, move on as easier questions might appear late on the paper.
3. Learn the proofs for some of the formulae – compared to other boards the CIE Formula sheet (MF 10) is not very useful at all. The Pure Mathematics section contains formulae that you'll rarely use (Power series expansions which are not in the syllabus) and the formulae you'll need are not in there. Thus, it's useful for you to ensure that you can do the problems when the formula is not in front of you.
4. Check your work after finishing every problem. Learn how to utilise your calculator for checking as some questions are tedious and have a lot of room for mistakes.
5. Ensure that you're familiar to work with constants. In FM we tend to use constants instead of actual figures for accuracy and generalizing purposes. Also use more significant figures (5 or 6) for accuracy purposes especially in the Statistics section.

DON'TS

1. Do not skip or underestimate any topic while studying this subject. Do keep in mind that there is only one question that is available as a choice while all other questions are compulsory. A weak foundation (as in little knowledge in the Maths papers) will also be a big obstacle in getting good results for this paper.
2. Do not read the contents unless you mastered the pre-requisite knowledge part of each Chapter. Most people skip M2 and S2 or do Paper 1 without a foundation in Maths P3 and ended up struggling with the notation and theories badly. Ensure that you have a strong grip with the pre-requisite contents before starting, as these set of notes if not sufficient if one is using them to prepare for M2 or S2 papers.
3. Do not rely solely on this set of notes. These notes are written as a reference guide to students and reading them without any form of practice will not get you an A* but will give you a C. It is also important to get problem sets from textbooks and work on past-year papers as those are the real keys to get good grades.
4. Do not rely too much on resources on other boards. Due to lack of resources for CIE Further Mathematics many candidates referred to textbooks and problems from other boards – but do keep in mind that not all topics in CIE are covered there. Most notably, other boards cover hyperbolic functions while CIE covers linear algebra.
5. Do not be discouraged by others' perspectives towards FM. It is not an easy subject and many chose not to take it up. High grade thresholds and the public's opinion may make FM look intimidating but hard work pays off in the end. An A* for Further Mathematics may seem hard but is possible if a lot of effort is put into it. Plus, the glory for acing in one of the toughest subjects in A-Levels will be rewarding.

4. Features of this set of notes:

The main things that differ it from other notes and normal course books

- **ALL IN ONE**

One major hurdle for CIE Further Mathematics students is that they lack an all-in-one course book, resulting in them receiving incomplete information as they search other boards for relevant concepts. In the process of writing this set of notes I referred to many books that are mentioned in the reference section and made sure that this set of notes completely covers every bit of content needed to master this subject. Still, I recommend that you check out these books as I did not include practice problems in this set of notes. Readers are recommended to get the books for more practice as the amount of past year problems available are limited.

- **CONCEPTS EXPLAINED THROUGH STORY AND PROOF**

One important skill for mastery in mathematics is that one is recommended to have a voice in their mind, telling them what to do when they are challenged by tough problems. For the purpose of achieving that, I included mind sets and procedures to tackle exam problems in my notes. Mind sets are used to help adapt the reader to the abstractness of higher-level mathematics while procedures to solve typical exam problems help the candidates to start doing so in the exams. Also, stories and proofs are used to explain the mathematical concepts in a less-formal manner compared to contents in university-level books as I attempt to introduce readers to the more formal method using this way.

- **PRE-REQUISITE KNOWLEDGE**

For the Pure Mathematics section, pre-requisite knowledge for Mathematics Paper 3 are briefly discussed before entering new contents in FM. For the Mechanics and Statistics, the same is done for Mathematics Paper 4 and Paper 6. This is done to refresh the reader's minds and is by no means a complete reference to those papers. Also for Paper 2 relevant contents from Mathematics Paper 5 and Paper 7 are briefly discussed with examples to ensure that readers are already exposed to the fundamentals before starting more challenging contents.

- **A RELIABLE RESOURCE**

This set of notes is a helpful resource to candidates attempting past year problems. Most examples I used to write the problems are CIE Past Year Problems are from years 2002 to 2007. The mark schemes for these papers are unavailable online, thus these notes can serve as good reference to anyone that wished to get some guidelines to those papers, as some problems are asked in a different manner than the more recent papers and are generally considered to be more challenging.

FURTHER MATHEMATICS PAPER 1 PURE MATHEMATICS

CHAPTER 1: POLYNOMIALS AND RATIONAL FUNCTIONS

1.1 Polynomial Equations

A) Roots of Polynomial Equations (Vieta's formula)

- LO: recall and use the relations between the roots and coefficients of polynomial equations, for equations of degree 2, 3, 4 only.

1. Quadratic Equations

When the equation $ax^2 + bx + c = 0$ has roots α and β:

$$(x - \alpha)(x - \beta) = 0 \rightarrow x^2 - (\alpha + \beta)x + \alpha\beta = 0.$$

Rewriting the equation into $x^2 + \frac{b}{a}x + \frac{c}{a} = 0$ and comparing coefficients:

Sum of roots:	$\alpha + \beta = -\frac{b}{a}$
Product of roots:	$\alpha\beta = \frac{c}{a}$

2. Cubic Equations

When the equation $ax^3 + bx^2 + cx + d = 0$ has roots α, β and γ:

$$(x - \alpha)(x - \beta)(x - \gamma) = 0 \rightarrow x^3 - (\alpha + \beta + \gamma)x^2 + (\alpha\beta + \beta\gamma + \gamma\alpha)x - \alpha\beta\gamma = 0$$

Rewriting the equation into $x^3 + \frac{b}{a}x^2 + \frac{c}{a}x + \frac{d}{a} = 0$ and comparing coefficients:

Sum of roots:	$\alpha + \beta + \gamma = -\frac{b}{a}$
Sum of pairs of roots:	$\alpha\beta + \beta\gamma + \gamma\alpha = \frac{c}{a}$
Product of roots:	$\alpha\beta\gamma = -\frac{d}{a}$

3. Quartic Equations

When the equation $ax^4 + bx^3 + cx^2 + dx + e = 0$ has roots α, β, γ and δ:

$$(x - \alpha)(x - \beta)(x - \gamma)(x - \delta) = 0$$
$$x^4 - (\alpha + \beta + \gamma + \delta)x^3 + (\alpha\beta + \beta\gamma + \gamma\alpha + \alpha\delta + \beta\delta + \gamma\delta)x^2$$
$$-(\alpha\beta\gamma + \alpha\beta\delta + \alpha\gamma\delta + \beta\gamma\delta) + \alpha\beta\gamma\delta = 0$$

Rewriting the equation into $x^4 + \frac{b}{a}x^3 + \frac{c}{a}x^2 + \frac{d}{a}x + \frac{e}{a} = 0$ and comparing coefficients:

Sum of roots:	$\alpha+\beta+\gamma+\delta=-\frac{b}{a}$
Sum of pairs of roots:	$\alpha\beta+\beta\gamma+\gamma\alpha+\alpha\delta+\delta+\gamma\delta=\frac{c}{a}$
Sum of triplets of roots:	$\alpha\beta\gamma+\alpha\beta\delta+\alpha\gamma\delta+\beta\gamma\delta=-\frac{d}{a}$
Product of roots:	$\alpha\beta\gamma\delta=\frac{e}{a}$

An easy trick to remember the identities:

When you add the roots one at a time: you get $-\frac{b}{a}$

When you add the roots two at a time: you get $\frac{c}{a}$

When you add the roots three at a time: you get $-\frac{d}{a}$

When you add the roots four at a time: you get $\frac{e}{a}$

Similarly: it's $-\frac{f}{a}$ for five roots at a time, $\frac{g}{a}$ for six roots at a time, and so on.

4. Common Identities used in problem solving:

- Complex-conjugate Root theorem: For any polynomial with real coefficients only, if $a + bi$ is a root of the polynomial, so is $a - bi$.
- For all $x \in R, x^2 \geq 0$.
- $(x+y)^2 \equiv x^2+2xy+y^2$
- $x^2+y^2+z^2 \equiv (x+y+z)^2-2(xy+yz+xz)$
- $w^2+x^2+y^2+z^2 \equiv (w+x+y+z)^2-2(wx+wy+wz+xy+yz+xz)$

Example 1: W2003 P1 Q6

Find the sum of the squares of the roots of the equation

$$x^3+x+12=0,$$

and deduce that only one of the roots is real. [4]

The real root of the equation is denoted by α. Prove that $-3 < \alpha < -2$, and hence prove that the modulus of each of the other roots lies between 2 and $\sqrt{6}$. [5]

$$\alpha^2+\beta^2+\gamma^2=(\alpha+\beta+\gamma)^2-2(\alpha\beta+\beta\gamma+\alpha\gamma)$$
$$=0^2-2(1)=-2$$

Since complex roots occur in conjugate pairs, a cubic equation only has one or 3 real roots.
But because sum of squares of roots is less than 0, not all roots are real.
Hence the cubic equation cannot have 3 real roots and thus only has one real root.
Let $f(x)=x^3+x+12$. We have $f(-3)=-18$ and $f(-2)=2$.

Since sign of $f(x)$ changes in the interval $-3 < x < -2$, there must exist a value α in that interval that satisfies $f(x) = 0$. Hence $f(x)$ has a root α where $-3 < \alpha < -2$.

Let the two complex roots be β and γ. By complex-root conjugate theorem, for real numbers p and q, we have $\beta = p + qi$ and $\gamma = p - qi$ and $|\beta| = |\gamma| = \sqrt{p^2 + q^2}$.
Product of roots: $\alpha\beta\gamma = \alpha(p + qi)(p - qi) = \alpha(p^2 + q^2) = -12 \rightarrow p^2 + q^2 = \frac{-12}{\alpha}$
Since $-3 < \alpha < -2$, $4 < -\frac{12}{\alpha} < 6 \rightarrow 4 < p^2 + q^2 < 6 \rightarrow 2 < \sqrt{p^2 + q^2} < \sqrt{6}$.
Hence the modulus of the two other roots are between 2 and $\sqrt{6}$.

Example 2: S2011 P13 Q3

Find a cubic equation with roots α, β and γ, given that

$$\alpha + \beta + \gamma = -6, \qquad \alpha^2 + \beta^2 + \gamma^2 = 38, \qquad \alpha\beta\gamma = 30. \qquad [3]$$

Hence find the numerical values of the roots. [3]

The cubic equation is $x^3 + ax^2 + bx + c = 0$, where

$$a = -(\alpha + \beta + \gamma) = 6$$

$$b = \alpha\beta + \beta\gamma + \gamma\alpha = \frac{1}{2}[(\alpha + \beta + \gamma)^2 - (\alpha^2 + \beta^2 + \gamma^2)] = \frac{1}{2}(36 - 38) = -1$$

$$c = -\alpha\beta\gamma = -30$$

Hence the equation is $x^3 + 6x^2 - x - 30 = 0$.
Using long division, $x^3 + 6x^2 - x - 30 = (x - 2)(x + 3)(x + 5)$
Hence α, β, γ are any permutations of $2, -3$ and -5.

Example 3: S2011 P11 Q2

The roots of the equation

$$x^3 + px^2 + qx + r = 0$$

are $\frac{\beta}{k}$, β, $k\beta$, where p, q, r, k and β are non-zero real constants. Show that $\beta = -\frac{q}{p}$. [4]

Deduce that $rp^3 = q^3$. [2]

Sum of roots: $\frac{\beta}{k} + \beta + k\beta = -p$

Sum of pairs of roots: $\frac{\beta^2}{k} + \beta^2 + k\beta^2 = q$

$$\beta\left(\frac{\beta}{k} + \beta + k\beta\right) = q$$

$$-p\beta = q \rightarrow \beta = -\frac{q}{p}$$

Product of roots: $-r = \beta^3 = -\frac{q^3}{p^3}$ $\rightarrow r = \frac{q^3}{p^3} \rightarrow q^3 = rp^3$

B) Other Techniques Used on Polynomials

- LO: use a given simple substitution to obtain an equation whose roots are related in a simple way to those of the original equation.

1. Substitutions

Consider the equation $x^3 - 6x^2 + 11x - 6 = 0$ with roots α, β and γ (1, 2 and 3 to be exact). Substitutions can be used to find equations with roots related to the original equation.

- An equation with roots α^2, β^2 and γ^2:

If α, β, γ satisfies $x^3 - 6x^2 + 11x - 6 = 0$, α^2, β^2 and γ^2 satisfies the new equation in *y*, where $y = x^2 \rightarrow x = \sqrt{y}$. Substitute that into the original equation:

$$y^{1.5} - 6y + 11y^{0.5} - 6 = 0 \rightarrow y^{0.5}(y + 11) = 6y + 6$$

Squaring both sides (notice how I removed the surds by grouping and factorising)

$$y(y^2 + 22y + 121) = 36y^2 + 72y + 36$$

$$y^3 + 22y^2 + 121y = 36y^2 + 72y + 36 \rightarrow y^3 - 14y^2 + 49y - 36 = 0$$

- An equation with roots $-\alpha + \beta + \gamma, \alpha - \beta + \gamma$and $\alpha + \beta - \gamma$:

Notice that $\alpha + \beta + \gamma = 6 \rightarrow \alpha + \beta - \gamma = 6 - 2\gamma$. Thus use the substitution $y = 6 - 2x$. Rearranging that to $x = \frac{1}{2}(6 - y)$:

$$\frac{1}{8}(6 - y)^3 - \frac{6}{4}(6 - y)^2 + \frac{11}{2}(6 - y) - 6 = 0$$

$$(6 - y)^3 - 12(6 - y)^2 + 44(6 - y) - 48 = 0$$

$$216 - 108y + 18y^2 - y^3 - 432 + 144y - 12y^2 + 216 - 44y = 0$$

$$-y^3 + 6y^2 - 8y = 0 \rightarrow y^3 - 6y^2 + 8y = 0$$

2. Master Polynomial Equation (learn the proof)

Consider the cubic equation $ax^3 + bx^2 + cx + d = 0$ with α, β and γ. Hence

$a\alpha^3 + b\alpha^2 + c\alpha + d = 0 \quad -(1) \quad (1) \times \alpha^n : (4) - \quad a\alpha^{n+3} + b\alpha^{n+2} + c\alpha^{n+1} + d\alpha^n = 0$

$a\beta^3 + b\beta^2 + c\beta + d = 0 \quad -(2) \quad (2) \times \beta^n : (5) - \quad a\beta^{n+3} + b\beta^{n+2} + c\beta^{n+1} + d\beta^n = 0$

$a\gamma^3 + b\gamma^2 + c\gamma + d = 0 \quad -(3) \quad (3) \times \gamma^n : (6) - \quad a\gamma^{n+3} + b\gamma^{n+2} + c\gamma^{n+1} + d\gamma^n = 0$

$(4) + (5) + (6)$: $a(\alpha^{n+3} + \beta^{n+3} + \gamma^{n+3}) + b(\alpha^{n+2} + \beta^{n+2} + \gamma^{n+2})$

$+c(\alpha^{n+1} + \beta^{n+1} + \gamma^{n+1}) + d(\alpha^n + \beta^n + \gamma^n) = 0$

Let $S_n = \alpha^n + \beta^n + \gamma^n$: $\boxed{aS_{n+3} + bS_{n+2} + cS_{n+1} + dS_n = 0}$

- This technique is a recursive formula to find powers of roots (for example S_3 can be found when S_0, S_1 and S_2 are found using normal methods.
- Note: The cubic equation version is shown here. Use this as a reference to work out the quartic equation version of this equation (the first part of Example 3 below)

Example 1: W2006 P1 Q6

The roots of the equation

$$x^3 + x + 1 = 0$$

are α, β, γ. Show that the equation whose roots are

$$\frac{4\alpha + 1}{\alpha + 1}, \quad \frac{4\beta + 1}{\beta + 1}, \quad \frac{4\gamma + 1}{\gamma + 1}$$

is of the form

$$y^3 + py + q = 0,$$

where the numbers p and q are to be determined. [5]

Hence find the value of

$$\left(\frac{4\alpha + 1}{\alpha + 1}\right)^n + \left(\frac{4\beta + 1}{\beta + 1}\right)^n + \left(\frac{4\gamma + 1}{\gamma + 1}\right)^n,$$

for $n = 2$ and for $n = 3$. [4]

First we have to decide the substitution to use.

$$y = \frac{4x + 1}{x + 1} = 4 - \frac{3}{x + 1} \rightarrow x + 1 = \frac{3}{4 - y} \rightarrow x = \frac{y - 1}{4 - y}$$

Then substitute.

$$x^3 + x + 1 = 0 \rightarrow \left(\frac{y-1}{4-y}\right)^3 + \frac{3}{4-y} = 0$$

$$(y-1)^3 + 3(4-y)^2 = 0$$

$$y^3 - 3y^2 + 3y - 1 + 3(16 - 8y + y^2) = 0$$

$$y^3 - 21y + 47 = 0$$

Hence $p = -21$ and $q = 47$.

$$S_n = \left(\frac{4\alpha + 1}{\alpha + 1}\right)^n + \left(\frac{4\beta + 1}{\beta + 1}\right)^n + \left(\frac{4\gamma + 1}{\gamma + 1}\right)^n$$

$$S_2 = 0^2 - 2(-21) = 42$$

Using the master equation, where $n = 0$:

$$S_3 = 21S_1 - 47S_0 = -47(3) = -141$$

Example 2: S2004 P1 Q11 OR

The roots of the equation

$$x^3 - x - 1 = 0$$

are α, β, γ, and

$$S_n = \alpha^n + \beta^n + \gamma^n.$$

(i) Use the relation $y = x^2$ to show that $\alpha^2, \beta^2, \gamma^2$ are the roots of the equation

$$y^3 - 2y^2 + y - 1 = 0.$$ [3]

(ii) Hence, or otherwise, find the value of S_4. [2]

(iii) Find the values of S_8, S_{12} and S_{16}. [9]

(i) $$x^3 - x - 1 = 0 \rightarrow y^{1.5} - y^{0.5} - 1 = 0 \rightarrow \sqrt{y}(y-1) = 1$$

Squaring both sides, $y(y^2 - 2y + 1) = 1 \rightarrow y^3 - 2y^2 + y - 1 = 0$

Since α, β, γ satisfies $x^3 - x - 1 = 0$, $\alpha^2, \beta^2, \gamma^2$ satisfies $y^3 - 2y^2 + y - 1 = 0$.

(ii) Using the equation $y^3 - 2y^2 + y - 1 = 0$,

$$\alpha^2 + \beta^2 + \gamma^2 = 2 \text{ and } \alpha^2\beta^2 + \gamma^2\alpha^2 + \beta^2\gamma^2 = 1.$$

Hence

$$S_4 = \alpha^4 + \beta^4 + \gamma^4 = \left(\alpha^2 + \beta^2 + \gamma^2\right)^2 - 2(\alpha^2\beta^2 + \gamma^2\alpha^2 + \beta^2\gamma^2) = 2^2 - 2 = 2$$

(iii) Similarly, let $z = y^2$, $\sqrt{z}(z+1) = 1 + 2z$.

Squaring both sides, $z(z^2 + 2z + 1) = 1 + 4z + 4z^2 \rightarrow z^3 - 2z^2 - 3z - 1 = 0$.

$$\alpha^4 + \beta^4 + \gamma^4 = 2 \text{ and } \alpha^4\beta^4 + \gamma^4\alpha^4 + \beta^4\gamma^4 = -3.$$

Hence $S_8 = \left(\alpha^4 + \beta^4 + \gamma^4\right)^2 - 2(\alpha^4\beta^4 + \gamma^4\alpha^4 + \beta^4\gamma^4) = 2^2 - 2(-3) = 10$.

From here we will use a modified equation of the master equation:

$$S_{4(n+3)} - 2S_{4(n+2)} - 3S_{4(n+1)} - S_{4n} = 0$$

When $n = 0$:

$$S_{12} = 2S_8 + 3S_4 + S_0 = 20 + 6 + 3 = 29$$

When $n = 1$:

$$S_{16} = 2S_{12} + 3S_8 + S_4 = 58 + 30 + 2 = 90$$

Example 3: S2008 P1 Q12 OR

The roots of the equation

$$x^4 - 5x^2 + 2x - 1 = 0$$

are $\alpha, \beta, \gamma, \delta$. Let $S_n = \alpha^n + \beta^n + \gamma^n + \delta^n$.

(i) Show that

$$S_{n+4} - 5S_{n+2} + 2S_{n+1} - S_n = 0. \quad [2]$$

(ii) Find the values of S_2 and S_4. [3]

(iii) Find the value of S_3 and hence find the value of S_6. [6]

(iv) Hence find the value of

$$\alpha^2(\beta^4 + \gamma^4 + \delta^4) + \beta^2(\gamma^4 + \delta^4 + \alpha^4) + \gamma^2(\delta^4 + \alpha^4 + \beta^4) + \delta^2(\alpha^4 + \beta^4 + \gamma^4). \quad [3]$$

(i) Refer to the notes

(ii) $S_2 = \alpha^2 + \beta^2 + \gamma^2 + \delta^2 = (\alpha + \beta + \gamma + \delta)^2 - 2(\alpha\beta + \alpha\gamma + \alpha\delta + \beta\gamma + \beta\delta + \gamma\delta)$

$S_2 = 0^2 - 2(-5) = 10$

Let $n = 0$: $S_0 = \alpha^0 + \beta^0 + \gamma^0 + \delta^0 = 4$, $S_1 = \alpha + \beta + \gamma + \delta = 0$

$$\begin{aligned} S_4 &= 5S_2 - 2S_1 + S_0 \\ &= 50 - 0 + 4 = 54 \end{aligned}$$

(iii) $$S_{-1} = \frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma} + \frac{1}{\delta} = \frac{\alpha\beta\gamma + \alpha\beta\delta + \alpha\gamma\delta + \beta\gamma\delta}{\alpha\beta\gamma\delta} = \frac{-2}{-1} = 2$$

Let $n = -1$: $$\begin{aligned} S_3 &= 5S_1 - 2S_0 + S_{-1} \\ &= 0 - 8 + 2 = -6 \end{aligned}$$

Let $n = 2$: $$\begin{aligned} S_6 &= 5S_4 - 2S_3 + S_2 \\ &= 270 + 12 + 10 = 292 \end{aligned}$$

(iv) $\alpha^2(\beta^4 + \gamma^4 + \delta^4) + \beta^2(\alpha^4 + \gamma^4 + \delta^4) + \gamma^2(\beta^4 + \alpha^4 + \delta^4) + \delta^2(\beta^4 + \gamma^4 + \alpha^4)$

$= \alpha^2(S_4 - \alpha^4) + \beta^2(S_4 - \beta^4) + \gamma^2(S_4 - \gamma^4) + \delta^2(S_4 - \delta^4)$

$= \alpha^2 S_4 + \beta^2 S_4 + \gamma^2 S_4 + \delta^2 S_4 - \alpha^6 - \beta^6 - \gamma^6 - \delta^6$

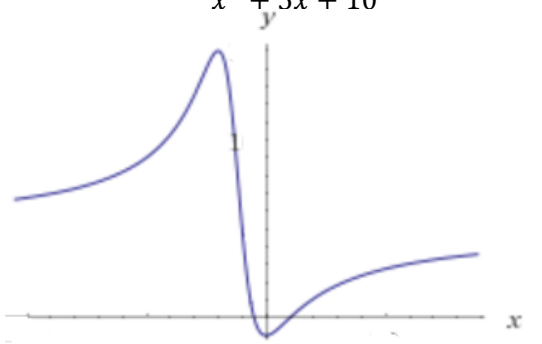
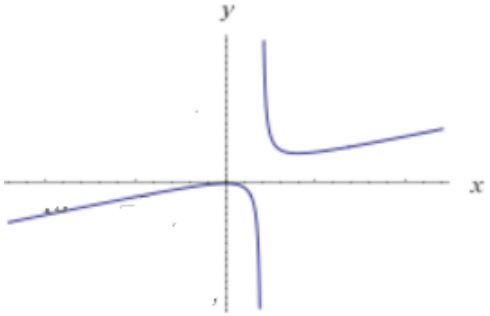
$= S_2 S_4 - S_6$

$= 10(54) - 292 = 248$

1.2 Sketching Rational Function Curves

- LO: sketch graphs of simple rational functions, including the determination of oblique asymptotes, in cases where the degree of the numerator and the denominator are at most 2 (detailed plotting of curves will not be required, but sketches will generally be expected to show significant features, such as the domain, range, turning points, asymptotes and intersections with the axes).

➢ A rational function can be expressed in the form $\frac{p(x)}{q(x)}$, where $p(x)$ and $q(x)$are polynomials. Some graphs of rational function curves are as shown below:

W2011 P13 Q10: $y = \frac{5(x^2 - x - 2)}{x^2 + 5x + 10}$	S2012 P13 Q6: $y = \frac{x^2}{x - 2}$
	

➢ The following curves are plotted using computer software. In the examination, you will only be requested to sketch graphs, but the significant features of the graph must be shown.
➢ Steps in sketching a rational function graph include the following:

Note: The following steps are presented as a guide; you will only be asked the steps as requested by the question (not all steps will be needed to obtain full marks).

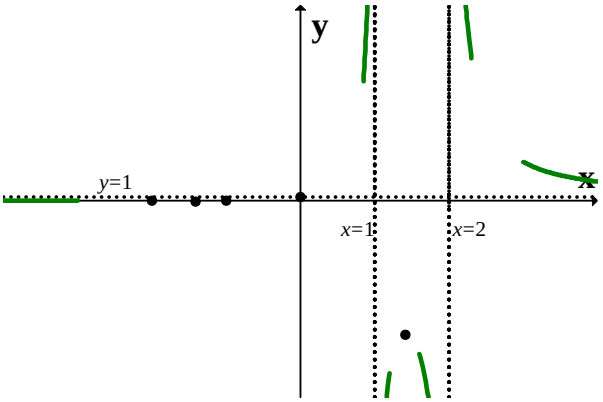
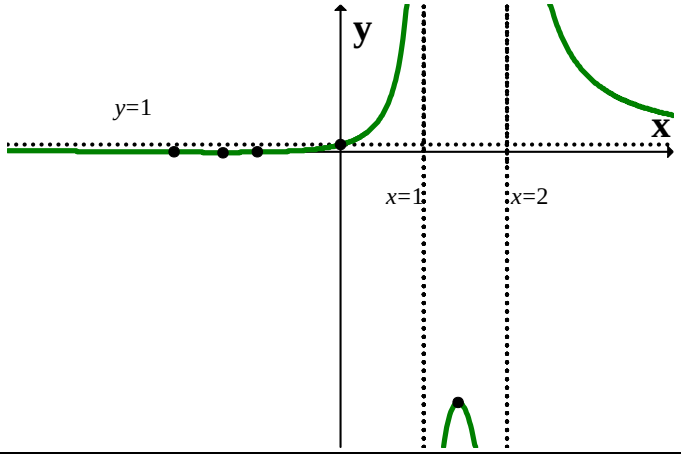
Definitions/ explanations are normally coloured, working is coloured in **bold**.

Example: Sketch the graph of

$$y = \frac{x^2 + 3x + 2}{(x - 1)(x - 2)}.$$

Step 1: Domain	➢ All the possible values that x can take. **x can take all values except 1 and 2, as a function cannot be divided by 0.** **The domain is $\{x: x \neq 1, x \neq 2, x \in R\}$**
Step 2: Range	➢ All the possible values that y can take. This is usually done by considering the discriminant of a quadratic equation.

	$$\boldsymbol{y=\frac{x^2+3x+2}{x^2-3x+2}}$$ $$\boldsymbol{x^2y-3xy+2y=x^2+3x+2}$$ $$\boldsymbol{x^2(y-1)-x(3y+3)+(2y-2)=0}$$ **If $x \in R$, then $B^2-4AC \geq 0$:** $$(3y-3)^2-4(y-1)(2y-2)\geq 0$$ $$9-18y+9y^2-8(1-2y+y^2)\geq 0$$ $$y^2-34y+1\geq 0$$ $$\boldsymbol{y\leq -17-12\sqrt{2},\quad y\geq -17+12\sqrt{2}}$$
Step 3: Intercepts	➢ There are two types of intercepts: ➢ *x*-intercepts: Values of x where $y=0$: Points where the curve touches the *x*-axis. **When $y=0$: $x^2+3x+2=0 \rightarrow (x+2)(x+1)=0$** **The *x*-intercepts are -1 and -2.** ➢ *y*-intercepts: Values of y where $x=0$: Points where the curve touches the *y*-axis. **When $x=0$: The *y*-intercept is** $$\boldsymbol{y=\frac{(0)^2+3(0)+2}{(0-1)(0-2)}=1}$$
Step 4: Turning Points	➢ If $\frac{dy}{dx}=0$ and $\frac{d^2y}{dx^2}>0$, the point is a minimum point. ➢ If $\frac{dy}{dx}=0$ and $\frac{d^2y}{dx^2}<0$, the point is a maximum point. $$\boldsymbol{y=\frac{x^2+3x+2}{x^2-3x+2}}$$ $$\boldsymbol{\frac{dy}{dx}=\frac{(x^2-3x+2)(2x+3)-(x^2+3x+2)(2x-3)}{(x^2-3x+2)^2}}$$ $$(x^2-3x+2)(2x+3)-(x^2+3x+2)(2x-3)=0$$ $$-6x^2+12=0 \rightarrow x=\pm\sqrt{2}$$ $$x=\pm\sqrt{2}, y=\frac{4\pm3\sqrt{2}}{4\mp3\sqrt{2}}\left(\frac{4\pm3\sqrt{2}}{4\pm3\sqrt{2}}\right)=\frac{16+18\pm24\sqrt{2}}{16-18}=-17\mp12\sqrt{2}$$ $$\boldsymbol{\frac{d^2y}{dx^2}=\frac{(x^2-3x+2)(-12x)+(6x^2-12)(4x-6)}{(x^2-3x+2)^3}}$$ $x=\sqrt{2}$: $\frac{d^2y}{dx^2}=\frac{(-12\sqrt{2})}{(4-3\sqrt{2})^2}<0 \rightarrow$ $(\sqrt{2},-17-12\sqrt{2})$ is a maximum point. $x=-\sqrt{2}$: $\frac{d^2y}{dx^2}=\frac{(12\sqrt{2})}{(4+3\sqrt{2})^2}>0 \rightarrow (-\sqrt{2},-17+12\sqrt{2})$ is a minimum point.

Step 5: Asymptotes	➢ A line or curve that describes the behaviour of the function at infinity. There are three types of asymptotes: 1. Vertical Asymptote If $x \to k$ as $y \to \infty$, then $x = k$ is a vertical asymptote. **When $x - 1 \to 0, y \to \infty$; When $x - 2 \to 0, y \to \infty$;** **The vertical asymptotes are $x = 1$ and $x = 2$.** 2. Horizontal Asymptote If $y \to h$ as $x \to \infty$, then $y = h$ is a horizontal asymptote. This occurs when the degree of the numerator equals the denominator. **Dividing by x^2:** $$y = \frac{1 + \frac{3}{x} + \frac{2}{x^2}}{(1 - \frac{1}{x})(1 - \frac{2}{x})}.$$ **As $x \to \infty$: $\frac{1}{x^n}$ for all positive integers n. $y \to \frac{1+0+0}{(1-0)(1-0)}$, $y \to 1$** **The horizontal asymptote is $y = 1$.** 3. Oblique Asymptote Oblique asymptotes only occur if the degree of the numerator exceeds the denominator. It can be obtained using long division, and will not exist if the graph already has horizontal asymptotes. **No oblique asymptotes here, as this graph has a horizontal asymptote (refer to examples below for oblique asymptotes)**
Step 6: Rough Sketch	➢ Begin by labelling the axes, intercepts and turning points. ➢ Draw the asymptotes and indicate the behaviour of the function there. 
Step 7: Full Sketch	➢ Use the above details to sketch the complete curve. 

- A sketch helps us to solve equations graphically. It also helps us to determine the behaviour of functions, especially when they are defined by other constants as seen below:

Example 2: W2002 P1 Q11 OR

The curve C has equation

$$y = \frac{(x-a)(x-b)}{x-c},$$

where a, b, c are constants, and it is given that $0 < a < b < c$.

(i) Express y in the form

$$x + P + \frac{Q}{x-c},$$

giving the constants P and Q in terms of a, b and c. [3]

(ii) Find the equations of the asymptotes of C. [2]

(iii) Show that C has two stationary points. [5]

(iv) Given also that $a + b > c$, sketch C, showing the asymptotes and the coordinates of the points of intersection of C with the axes. [4]

(i)

By long division: $\frac{x^2 - ax - bx + ab}{x-c} = x + (c-a-b) + \frac{c(c-a-b)+ab}{x-c}$

Comparing the terms: $P = c - a - b,\ Q = c(c-a-b) + ab = (c-a)(c-b)$

(ii)

$$y = x + (c-a-b) + \frac{(c-a)(c-b)}{x-c}$$

When $y \to \infty, x - c \to 0$; Vertical asymptote is $x = c$.

When $x \to \infty, \frac{Q}{x-c} \to 0; y \to x + P$; Oblique asymptote is $y = x + (c-a-b)$.

(iii)

$$y = x + P + Q(x-c)^{-1} \to \frac{dy}{dx} = 1 - Q(x-c)^{-2}$$

When $\frac{dy}{dx} = 0$: $(x-c)^2 = Q \to x = \pm\sqrt{Q} + c$

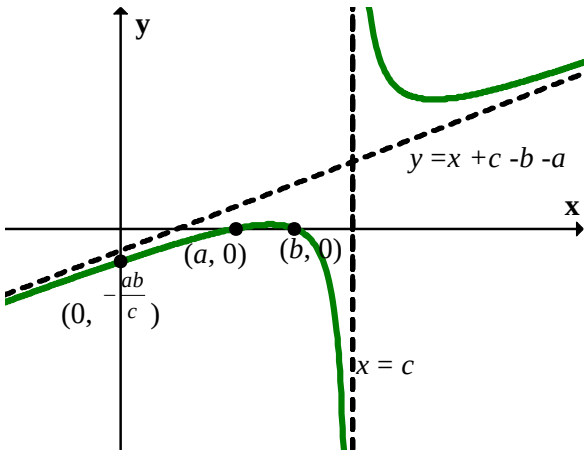
Since $Q > 0$, C has two stationary points at $x = \pm\sqrt{Q} + c$.

(iv) When $a + b > c, P < 0$.

The intercepts are

$(a, 0), (b, 0)$ *and* $\left(0, -\frac{ab}{c}\right)$.

A sketch of C is as follows:



Example 3: W2006 P1 Q12 EITHER

The curve C has equation

$$y = \frac{x^2 + qx + 1}{2x + 3},$$

where q is a positive constant.

(i) Obtain the equations of the asymptotes of C. [3]

(ii) Find the value of q for which the x-axis is a tangent to C, and sketch C in this case. [4]

(iii) Sketch C for the case $q = 3$, giving the exact coordinates of the points of intersection of C with the x-axis. [3]

(iv) It is given that, for all values of the constant λ, the line

$$y = \lambda x + \tfrac{3}{2}\lambda + \tfrac{1}{2}(q - 3)$$

passes through the point of intersection of the asymptotes of C. Use this result, with the diagrams you have drawn, to show that if $\lambda < \frac{1}{2}$ then the equation

$$\frac{x^2 + qx + 1}{2x + 3} = \lambda x + \tfrac{3}{2}\lambda + \tfrac{1}{2}(q - 3)$$

has no real solution if q has the value found in part **(ii)**, but has 2 real distinct solutions if $q = 3$. [4]

(i) Using long division, $y = \frac{1}{2}x + \frac{1}{2}\left(q - \frac{3}{2}\right) - \frac{6q+13}{4(2x+3)}$

When $y \to \infty, 2x + 3 \to 0$; Vertical asymptote is $x = -\frac{3}{2}$.

When $x \to \infty, \frac{1}{2x+3} \to 0; y \to \frac{1}{2}\left(x + q - \frac{3}{2}\right)$;

Oblique asymptote is $y = \frac{1}{2}\left(x + q - \frac{3}{2}\right)$.

(ii) When C intersects with x-axis:

$y = 0 \to x^2 + qx + 1 = 0$

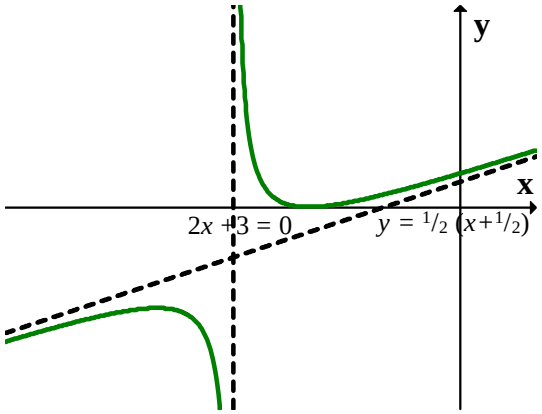
Since the x-axis is a tangent to C:

$$B^2 - 4AC = 0$$
$$q^2 - 4 = 0$$
$$q = 2, \text{since } q > 0.$$

Intercepts are $x = -1$ and $y = \frac{1}{3}$.

Oblique asymptote is $y = \frac{1}{2}\left(x + \frac{1}{2}\right)$.

A sketch of C is as follows:

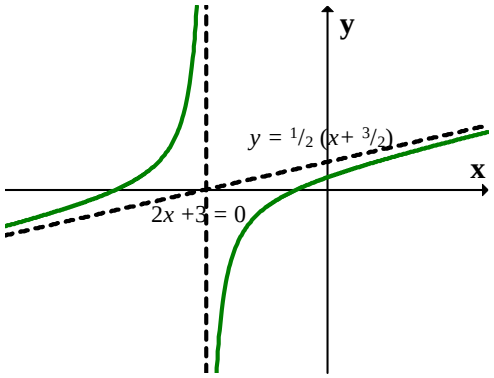


(iii) When $q = 3, y = 0$, x-intercepts are:

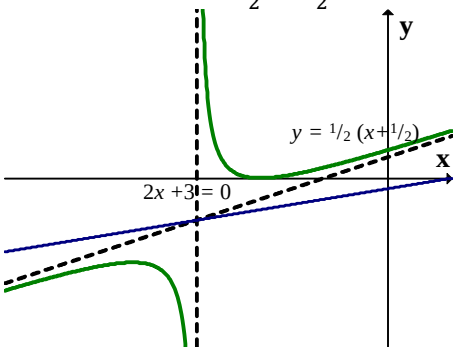
$$x^2 + 3x + 1 = 0 \to x = \frac{-3 \pm \sqrt{9 - 4}}{2} = \frac{-3 \pm \sqrt{5}}{2}$$

Oblique asymptote is $y = \frac{1}{2}\left(x + \frac{3}{2}\right)$. y-intercept is still $\frac{1}{3}$.

A sketch of C is as follows:

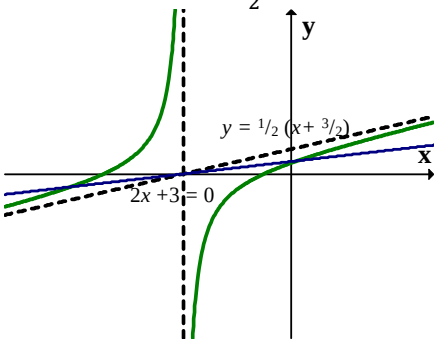


(iv) When $q = 2$:
the line is $y = \lambda x + \frac{3}{2}\lambda - \frac{1}{2}$.



0 intersecting points, hence 0 real roots.

When $q = 3$:
the line is $y = \lambda x + \frac{3}{2}\lambda$.



2 intersecting points, hence 2 real roots.

Example 4: S2007 P1 Q12 EITHER

The curve C has equation

$$y = \lambda x + \frac{x}{x+2},$$

where λ is a non-zero constant.

(i) Find the asymptotes of C. [3]

(ii) Show that if $\lambda > 0$ then $\frac{dy}{dx} > 0$ at all points of C. [2]

(iii) Show that, for $\lambda < -\frac{1}{2}$, C has two distinct stationary points, both to the left of the y-axis. [3]

(iv) In separate diagrams draw sketches of C for each of the cases $\lambda > 0$ and $\lambda < -\frac{1}{2}$. [6]

(i) When $y \to \infty, x + 2 \to 0$; Vertical asymptote is $x = -2$.

$$y = \lambda x + 1 - \frac{2}{x+2}$$

When $x \to \infty, \frac{1}{x+2} \to 0; y \to \lambda x + 1$; Oblique asymptote is $y = \lambda x + 1$.

(ii)
$$\frac{dy}{dx} = \lambda + 2(x+2)^{-2} = \lambda + \frac{2}{(x+2)^2}$$

For all real values x, $(x+2)^2 \geq 0$, thus if $\lambda > 0$, $\frac{dy}{dx} > 0$ for all points on C.

(iii) When $\frac{dy}{dx} = 0, (x+2)^2 = -\frac{2}{\lambda} \to x = -2 \pm \sqrt{-\frac{2}{\lambda}}$.

If $\lambda < -\frac{1}{2}, -\frac{2}{\lambda} < 4 \to \sqrt{-\frac{2}{\lambda}} < 2$.

It is obvious that $-2-\sqrt{-\frac{2}{\lambda}}<0$.

For $x+2=\sqrt{-\frac{2}{\lambda}}$, since $x+2<2, x<0$.

Since both turning points of C satisfies $x<0$, C has two turning points both to the left of the *x*-axis.

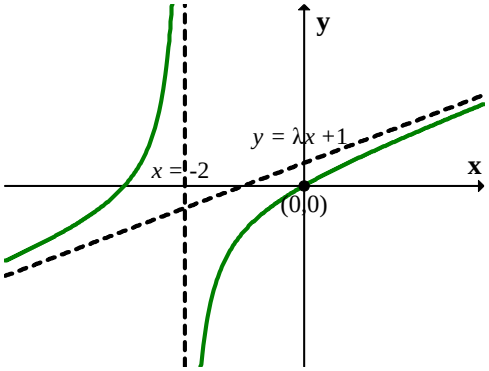
(iv) An intercept is (0, 0) in both cases.

For $\lambda>0$:

No turning points as $\frac{dy}{dx}>0$.

Oblique asymptote slants upwards.

A sketch of C is as follows:

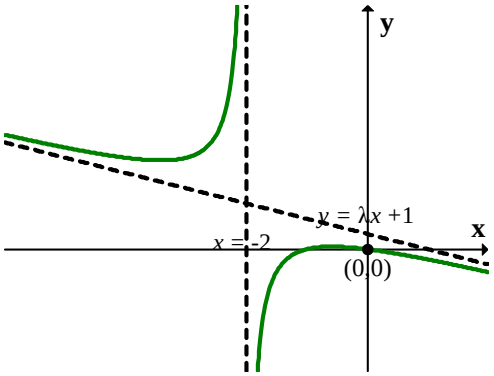


For $\lambda<-\frac{1}{2}$:

Two turning points to the left of *x*-axis.

Oblique asymptote slants downwards.

A sketch of C is as follows:



Example 5: W2008 P1 Q12 EITHER

The curve C has equation

$$y=\frac{(x-2)(x-a)}{(x-1)(x-3)},$$

where a is a constant not equal to 1, 2 or 3.

(i) Write down the equations of the asymptotes of C. [2]

(ii) Show that C meets the asymptote parallel to the x-axis at the point where $x=\frac{2a-3}{a-2}$. [2]

(iii) Show that the x-coordinates of any stationary points on C satisfy

$$(a-2)x^2+(6-4a)x+(5a-6)=0,$$

and hence find the set of values of a for which C has stationary points. [6]

(iv) Sketch the graph of C for

(a) $a>3$,

(b) $2<a<3$.

[4]

(i) When $y\to\infty, x-1\to 0, x-3\to 0$; Vertical asymptotes are $x=1$ and $x=3$.

Dividing by x^2:

$$y=\frac{\left(1-\frac{2}{x}\right)\left(1-\frac{a}{x}\right)}{\left(1-\frac{1}{x}\right)\left(1-\frac{3}{x}\right)}$$

When $x\to\infty, \frac{n}{x}\to 0, y\to\frac{(1-0)(1-0)}{(1-0)(1-0)}=1$; Horizontal asymptote is $y=1$.

(ii) When x meets the horizontal asymptote:

$$\frac{(x-2)(x-a)}{(x-1)(x-3)}=1$$

$$x^2-2x-ax+2a=x^2-4x+3$$

$$2x-ax=3-2a \rightarrow x=\frac{3-2a}{2-a}=\frac{2a-3}{a-2}$$

(iii) $$\frac{dy}{dx}=\frac{(x-1)(x-3)(2x-2-a)-(x-2)(x-a)(2x-4)}{(x-1)^2(x-3)^2}$$

When $\frac{dy}{dx}=0$:

$$(x-1)(x-3)(2x-2-a)-(x-2)(x-a)(2x-4)=0$$

$$(x^2-4x+3)(2x-2-a)-(x-a)(2x^2-8x+8)=0$$

$$6x-2x^2-6-ax^2+4ax-3a+2ax^2-8ax+8a=0$$

$$(a-2)x^2+(6-4a)x+(5a-6)=0$$

To have stationary points, $x \epsilon R$, thus $B^2-4AC \geq 0$.

$$(6-4a)^2-4(a-2)(5a-6)\geq 0$$

$$9-12a+4a^2-(5a^2-16a+12)\geq 0$$

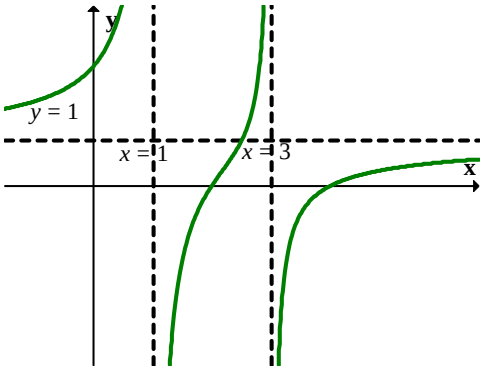
$$a^2-4a+3\leq 0$$

$$(a-3)(a-1)\leq 0 \rightarrow 1\leq a\leq 3$$

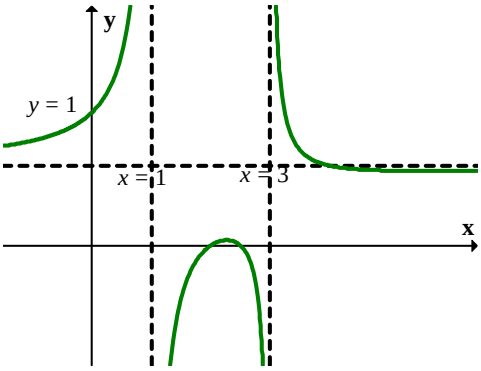
Since $a \neq 1, 2, 3$, the set of values is $\{a:\ 1<a<2, 2<a<3, a\ \in R\}$.

(iv) For both cases: x-intercepts are 2 and a, y-intercept is $\frac{2}{3}a$.

(a) For $a>3$:
curve meets the asymptote in the region $2<x<3$.
2nd x-intercept is greater than 3.
No turning points.
A sketch of C is as follows:



(b) For $2<a<3$:
curve meets the asymptote in the region $x>3$.
2nd x-intercept is between 2 and 3.
Two turning points.
A sketch of C is as follows:

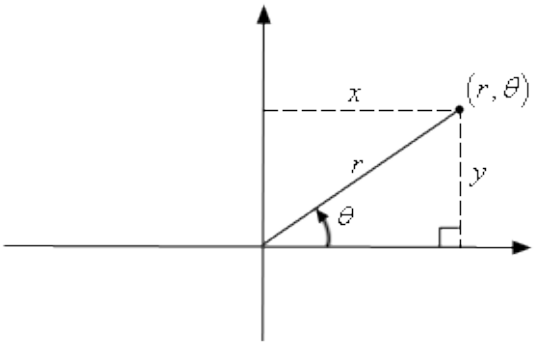


CHAPTER 2: POLAR COORDINATES

2.1 Cartesian and Polar Coordinates

- LO: understand the relations between cartesian and polar coordinates (using the convention $r \geq 0$), and convert equations of curves from cartesian to polar form and vice versa;
- LO: sketch simple polar curves, for $0 \leq \theta < 2\pi$ or $-\pi < \theta \leq \pi$ or a subset of these intervals, and identify significant features of polar curves such as symmetry, least/greatest values of r, and the form of the curve at the pole (knowledge that values of θ for which $r = 0$ give directions of tangents at the pole is included);

A) Conversions



Cartesian coordinates define a point in the *xy*-plane using coordinates (x, y), where: ➢ x represents the horizontal distance of the point from the *x*-axis, ➢ y represents the vertical distance of the point from the *y*-axis,	Polar coordinates defines a point in the *xy*-plane using coordinates (r, θ), where: ➢ r represents the distance from the point from the origin, ➢ θ represents the angle between the line made by the origin and the point and the positive *x*-axis.

➢ Some terminologies used in polar coordinates:

The origin (0, 0) is known as the **pole** ($\boldsymbol{r = 0}$):

The positive *x*-axis is known as the **initial line** ($\boldsymbol{\theta = 0}$).

➢ A cartesian equation is sometimes simplified when converted into polar form, using the identities below:

$$\boldsymbol{x = r\cos\theta - (1)} \qquad \boldsymbol{y = r\sin\theta - (2)}$$

$$\boldsymbol{(1)^2 + (2)^2 : r^2 = x^2 + y^2 \rightarrow r = \sqrt{x^2 + y^2}}$$

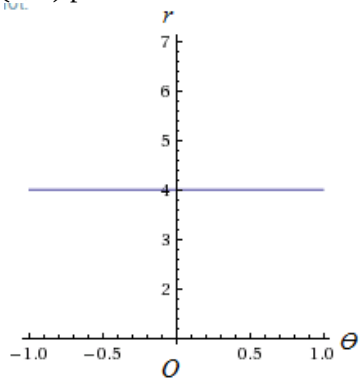
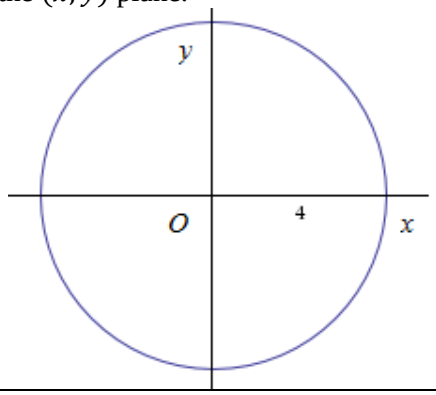
$$\boldsymbol{\frac{(2)}{(1)} : \tan\theta = \frac{y}{x} \rightarrow \theta = \tan^{-1}\left(\frac{y}{x}\right)}$$

- A few examples on how to do so is shown below:

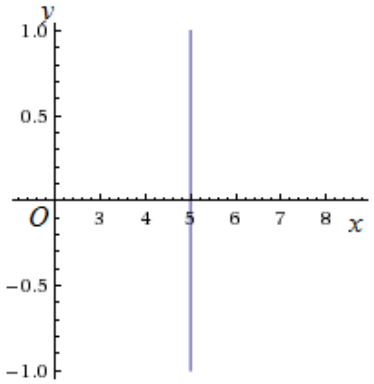
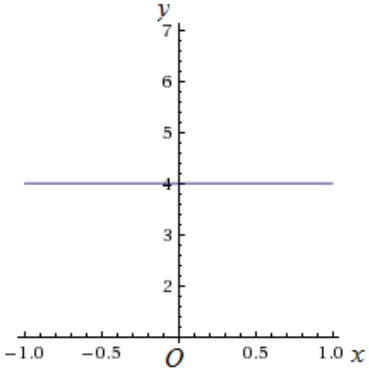
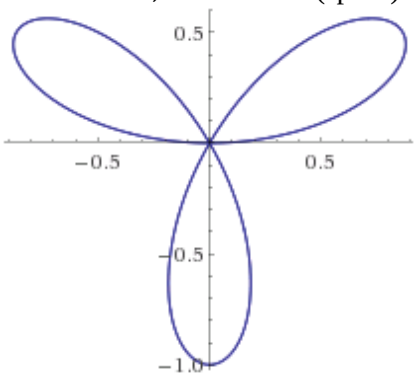
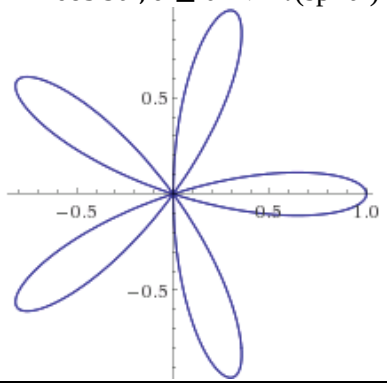
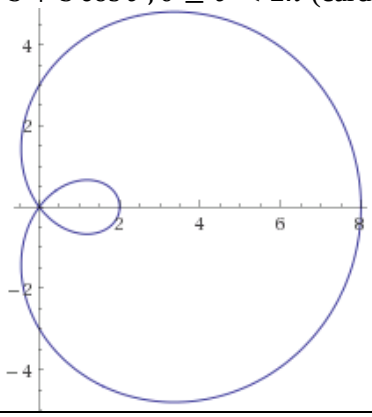
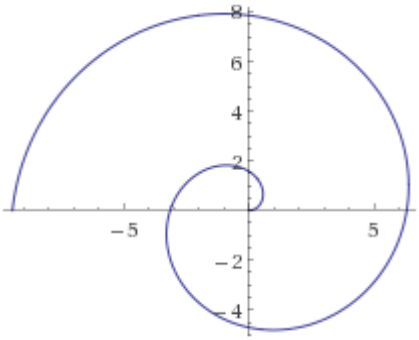
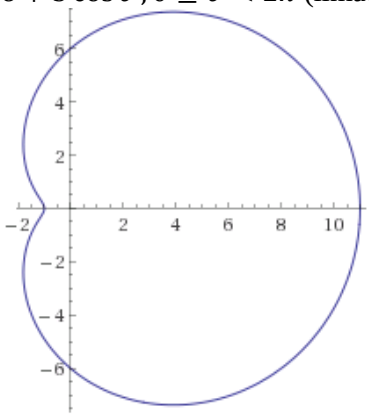
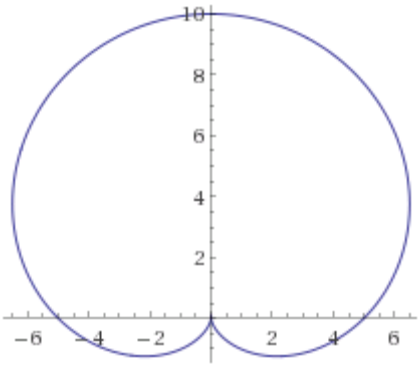
Converting from Cartesian to polar form:	Converting from polar to Cartesian form:
(5, 1) $r = \sqrt{5^2 + 1^2} = \sqrt{26}$ $\theta = \tan^{-1}\left(\frac{1}{5}\right) = 0.197\ rad$ Hence $(5, 1) \rightarrow \left(\sqrt{26}, 0.197\right)$.	(5, 1): $x = 5\cos 1 = 2.702$ $y = 5\sin 1 = 4.207$ Hence $(5, 1) \rightarrow (2.702, 4.207)$. (Notice that θ is always in radians and not degrees)
$(x^2 + y^2)^2 = x^2 - y^2$ $(r^2)^2 = r^2cos^2\theta - r^2sin^2\theta$ $r^4 = r^2(cos^2\theta - sin^2\theta) = r^2\cos 2\theta$ $r^2 = \cos 2\theta$ $r = \sqrt{\cos 2\theta}$ as $r \geq 0$.	$r = \sqrt{2}\sec\left(\theta - \frac{1}{4}\pi\right)$ $r\cos\left(\theta - \frac{1}{4}\pi\right) = \sqrt{2}$ $r\cos\theta\cos\frac{1}{4}\pi + r\sin\theta\sin\frac{1}{4}\pi = \sqrt{2}$ $\frac{x}{\sqrt{2}} + \frac{y}{\sqrt{2}} = \sqrt{2} \rightarrow x + y = 2$

B) Sketching Polar Graphs

- Polar graphs are sketched in the (x, y) plane instead of the (r, θ) plane. This causes the graphs to be very different as seen below:
- As seem below the function $r = 4$ is a horizontal line in the (r, θ) plane but is a circle in the (x, y) plane.

$r = 4$	
in the (r, θ) plane:	in the (x, y) plane:
	

➢ Sketches of a few curves in polar form are shown as below:

$r\cos\theta = 5$ (vertical line)	$r\sin\theta = 4$ (horizontal line)
	
$r = \sin 3\theta\,, 0 \le \theta < 2\pi$ (spiral)	$r = \cos 5\theta\,, 0 \le \theta < 2\pi$(spiral)
	
$r = 3 + 5\cos\theta\,, 0 \le \theta < 2\pi$ (cardioid)	$r = \theta, 0 \le \theta < 3\pi$ (spiral)
	
$r = 6 + 5\cos\theta\,, 0 \le \theta < 2\pi$ (limacon)	$r = 5 + 5\sin\theta\,, 0 \le \theta < 2\pi$ (cardioid)
	

- Some common forms of polar graphs that you should remember:

1. $\boldsymbol{r\cos\theta = a} \rightarrow x = a$ is a vertical line.

2. $\boldsymbol{r\sin\theta = a} \rightarrow y = a$ is a horizontal line.

3. $\boldsymbol{r = a} \rightarrow x^2 + y^2 = a^2$ is a circle with radius *a* and centre at the origin.

- For writing purposes my curves are accurate because they are drawn using a computer software, but one sketched in the exam hall just needs to show significant shapes and features.
- Some tips on sketching good polar curves:

1. Start by labelling the points on the intercepts (where $\theta = 0, \pm\frac{1}{2}\pi, \pm\pi$) in the axes as these are points where your curve must pass through and hence they are good guidelines.

2. Find out the points where the curve goes back to the origin (where$r = 0$) in your sketch.

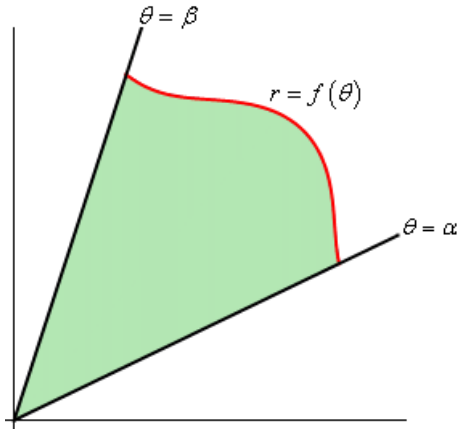
3. Recall the way the value of the function increase / decrease in certain intervals. This allows you to determine whether the point must be made further away or closer to the pole. For instance, since the value of $r = \sin\theta + 1$ decreases in the interval $\left[\frac{\pi}{2}, \pi\right]$, a sketch of the function must show that the graph gets closer to the pole in that interval.

4. Sometimes reducing the curve to Cartesian form helps you to sketch the curve (especially if it's one of the three forms mentioned above)

5. It is also useful if you are able to detect lines of symmetry (so that you can reflect the curve to produce an identical one on the other side of the graph when needed)

2.2 Area of a Sector in Polar Coordinates

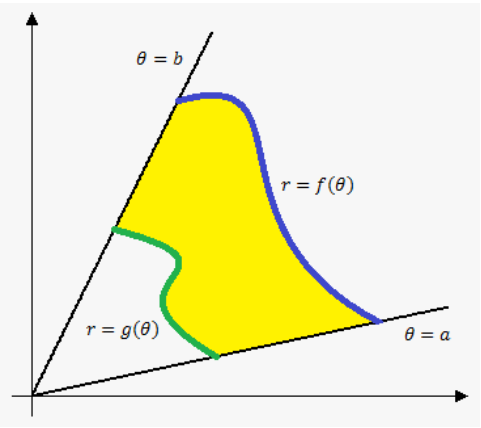
- LO: use the formula $\int_\alpha^\beta \tfrac{1}{2}r^2\,d\theta$ for the area of a sector in simple cases.



- Divide the angle between lines $\theta = \alpha$ and $\theta = \beta$ into many small angles of $\delta\theta$ each, we get many sectors with radius r and angle $\delta\theta$, as an arc is always a part of a circle.
- Recall from P1 Circular Measure that the area of a sector is $\frac{1}{2}r^2\theta$. Area of the shaded region, A is then obtained by summing all areas, hence

$$A = \sum \frac{1}{2}r^2\delta\theta = \int_\alpha^\beta \frac{1}{2}r^2 d\theta\,,$$

- as definite integration is defined as the summation of infinitely small areas.

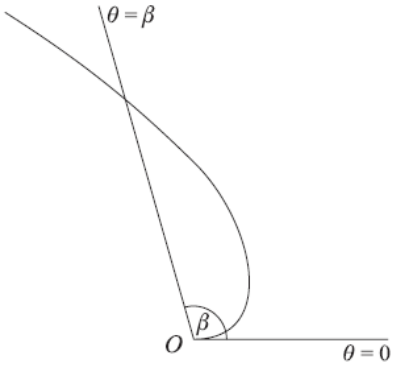


- For the area between two curves, the same applies as in the Cartesian case, hence area of yellow region, A:

$$A = \int_\alpha^\beta \frac{1}{2}\left[f(\theta)\right]^2 - \left[g(\theta)\right]^2 d\theta$$

- Reminder: This formula calculates area that is swept from the origin.
- The usual techniques for integration apply here. It is common here to use the double angle formula for cosine to integrate when r involves a function of $\sin\theta$ or $\cos\theta$.

Example 1: W2003 P1 Q1



The curve C has polar equation $r = \theta^{\frac{1}{2}}e^{\theta^2/\pi}$,

where $0 \leqslant \theta \leqslant \pi$. The area of the finite region bounded by C and the line $\theta = \beta$ is π (see diagram). Show that

$$\beta = (\pi \ln 3)^{\frac{1}{2}}. \qquad [6]$$

$$Area = \frac{1}{2}\int_0^{\beta} \theta e^{\frac{2}{\pi}\theta^2}\, d\theta$$

$$= \frac{1}{2}\left(\frac{\pi}{4}\right)\int_0^{\beta} \frac{4}{\pi}\theta e^{\frac{2}{\pi}\theta^2}\, d\theta$$

$$\frac{\pi}{8}\left[e^{\frac{2}{\pi}\theta^2}\right]_0^{\beta} = \pi$$

$$e^{\frac{2}{\pi}\beta^2} - e^0 = 8$$

$$\frac{2}{\pi}\beta^2 = \ln 9$$

$$= 2\ln 3$$

$$\beta^2 = \pi \ln 3$$

$$\beta = (\pi \ln 3)^{\frac{1}{2}}$$

Note: $\int f'(x)e^{f(x)}dx = e^{f(x)} + c$, by using the substitution $u = f(x)$.

Problem Solving:

1. General problems involve sketching curves, converting equations and calculating area.

2. For maximum or minimum distances, set $\frac{dr}{d\theta} = 0$ for pole and $\frac{dy}{d\theta} = 0$ for initial line.

3. Usual tests for maximum and minimum are used again here (sign of $\frac{d^2r}{d\theta^2} / \frac{d^2y}{d\theta^2}$).

Example 2: S2002 P1 Q11 OR

The curve C has polar equation

$$r = a(1 + \cos\theta), \qquad -\pi < \theta \leqslant \pi,$$

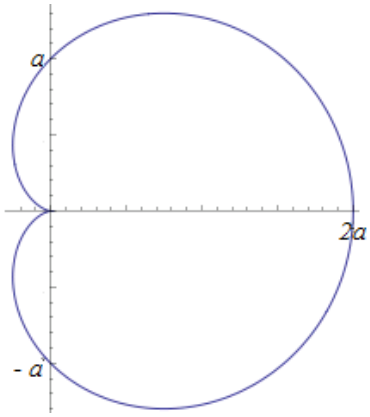
where a is a positive constant.

(i) Sketch C. [2]

(ii) Show that the area of the region enclosed by C is $\frac{3}{2}\pi a^2$. [5]

(iii) The point on C with polar coordinates (r, θ) has cartesian coordinates (x, y). Find the minimum value of y. [6]

(i)



(ii)

$$Area = \frac{1}{2}\int_{-\pi}^{\pi} a^2(1+\cos\theta)^2 d\theta$$

$$= \frac{1}{2}a^2\int_{-\pi}^{\pi} 1 + 2\cos\theta + \cos^2\theta d\theta$$

$$= \frac{1}{2}a^2\int_{-\pi}^{\pi} \frac{3}{2} + 2\cos\theta + \frac{1}{2}\cos 2\theta d\theta$$

$$= \frac{1}{2}a^2\left[\frac{3}{2}\theta + 2\sin\theta + \frac{1}{4}\sin 2\theta\right]_{-\pi}^{\pi}$$

$$= \frac{1}{8}a^2\left[6\theta + 8\sin\theta + \sin 2\theta\right]_{-\pi}^{\pi}$$

$$= \frac{1}{8}a^2\left[6\pi - (-6\pi)\right] = \frac{3}{2}\pi a^2$$

(iii)

$$y = r\sin\theta = a\sin\theta\,(1 + \cos\theta)$$

$$\frac{dy}{d\theta} = a\cos\theta\,(1 + \cos\theta) - a\sin\theta\,(\sin\theta)$$

$$\frac{dy}{d\theta} = 0, \qquad a(\cos\theta + cos^2\theta - sin^2\theta) = 0$$

$$2cos^2\theta + \cos\theta - 1 = 0$$

$$(2\cos\theta - 1)(\cos\theta + 1) = 0$$

$$\cos\theta = -1 \text{ or } \cos\theta = \frac{1}{2}$$

$$\frac{d^2y}{dx^2} = a\frac{d}{dx}(2cos^2\theta + \cos\theta - 1) = -a(4\sin\theta\cos\theta + \sin\theta)$$

When $\cos\theta = -1$, $\sin\theta = 0$, $\frac{d^2y}{dx^2} = 0$, hence the test fails.

When $\cos\theta = \frac{1}{2}$, $\sin\theta = \pm\frac{\sqrt{3}}{2}$:

$$\frac{d^2y}{dx^2} = \mp a\frac{\sqrt{3}}{2}\left(4\left(\frac{1}{2}\right) + 1\right) = \mp\frac{3\sqrt{3}}{2}$$

hence when $\cos\theta = \frac{1}{2}$, $\sin\theta = -\frac{\sqrt{3}}{2}$, $\frac{d^2y}{dx^2} = -\frac{3\sqrt{3}}{2} < 0$.

$$y_{min} = a\left(-\frac{\sqrt{3}}{2}\right)\left(1 + \frac{1}{2}\right) = -\frac{3\sqrt{3}}{4}a$$

Example 2: W2005 P1 Q12 OR

The curves C_1 and C_2 have polar equations

$$r = 4\cos\theta \quad \text{and} \quad r = 1 + \cos\theta$$

respectively, where $-\frac{1}{2}\pi \leqslant \theta \leqslant \frac{1}{2}\pi$.

(i) Show that C_1 and C_2 meet at the points $A\left(\frac{4}{3}, \alpha\right)$ and $B\left(\frac{4}{3}, -\alpha\right)$, where α is the acute angle such that $\cos\alpha = \frac{1}{3}$. [2]

(ii) In a single diagram, draw sketch graphs of C_1 and C_2. [3]

(iii) Show that the area of the region bounded by the arcs OA and OB of C_1, and the arc AB of C_2, is

$$4\pi - \tfrac{1}{3}\sqrt{2} - \tfrac{13}{2}\alpha.$$ [7]

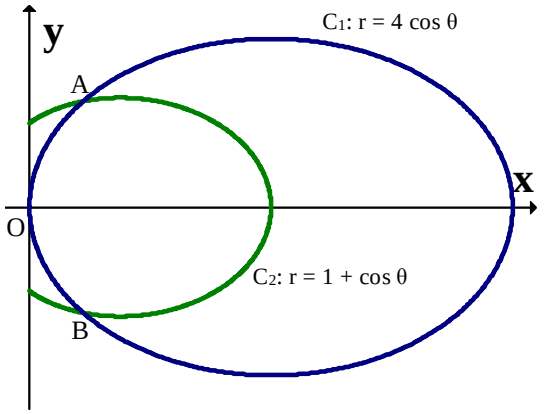
(i)

$$4\cos\theta = 1 + \cos\theta$$
$$3\cos\theta = 1$$
$$\cos\theta = \frac{1}{3}$$
$$\theta = \pm\cos^{-1}\frac{1}{3}$$
$$r = 4\left(\frac{1}{3}\right) = \frac{4}{3}$$

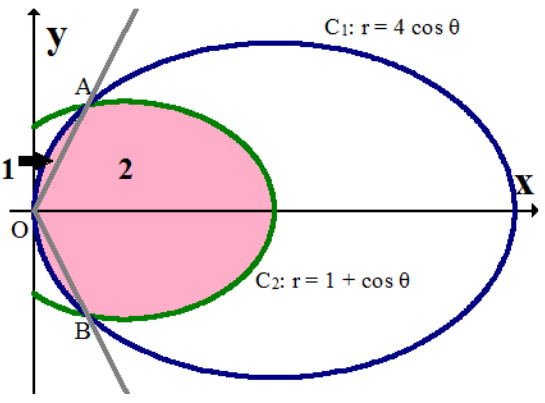
Intersecting points are hence

$$\left(\frac{4}{3}, \pm\cos^{-1}\frac{1}{3}\right) = \left(\frac{4}{3}, \pm\alpha\right)$$

(ii)



(iii)



$$\cos\alpha = \frac{1}{3}$$

$$\sin\alpha = \sqrt{1 - \frac{1}{9}} = \frac{2}{3}\sqrt{2}$$

$$\sin 2\alpha = 2\sin\alpha\cos\alpha = \frac{4}{9}\sqrt{2}$$

Area of region 1:

$$\frac{1}{2}\int_{\alpha}^{\frac{1}{2}\pi}(4\cos\theta)^2 d\theta$$
$$=\int_{\alpha}^{\frac{1}{2}\pi}8\cos^2\theta d\theta$$
$$=\int_{\alpha}^{\frac{1}{2}\pi}4+4\cos 2\theta d\theta$$
$$=\left[4\theta+2\sin 2\theta\right]_{\alpha}^{\frac{1}{2}\pi}$$
$$=2\pi-4\alpha-\frac{8}{9}\sqrt{2}$$

Area of region 2:

$$\frac{1}{2}\int_0^{\alpha}(1+\cos\theta)^2 d\theta=\frac{1}{2}\int_0^{\alpha}1+2\cos\theta+\cos^2\theta d\theta$$
$$=\frac{1}{2}\int_0^{\alpha}\frac{3}{2}+2\cos\theta+\frac{1}{2}\cos 2\theta d\theta$$
$$=\frac{1}{2}\left[\frac{3}{2}\theta+2\sin\theta+\frac{1}{4}\sin 2\theta\right]_0^{\alpha}$$
$$=\frac{1}{8}\left[6+8\sin\theta+\sin 2\theta\right]_0^{\alpha}$$
$$=\frac{1}{8}\left(6\alpha+\frac{16}{3}\sqrt{2}+\frac{4}{9}\sqrt{2}\right)$$
$$=\frac{3}{4}\alpha+\frac{13}{18}\sqrt{2}$$

Area of shaded region $= 2$(Area of region 1 + Area of region 2)

$$=4\pi-8\alpha-\frac{16}{9}\sqrt{2}+\frac{3}{2}\alpha+\frac{13}{9}\sqrt{2}$$
$$=4\pi-\frac{13}{2}\alpha-\frac{1}{3}\sqrt{2}\text{ (SHOWN)}$$

Example 4: W2009 P1 Q10

The curve C has polar equation

$$r = a\sin 3\theta,$$

where $0 \leqslant \theta \leqslant \frac{1}{3}\pi$.

(i) Show that the area of the region enclosed by C is $\frac{1}{12}\pi a^2$. [3]

(ii) Show that, at the point of C at maximum distance from the initial line,

$$\tan 3\theta + 3\tan\theta = 0.$$ [3]

(iii) Use the formula

$$\tan 3\theta = \frac{3\tan\theta - \tan^3\theta}{1 - 3\tan^2\theta}$$

to find this maximum distance. [4]

(iv) Draw a sketch of C. [2]

(i)

$$Area=\frac{1}{2}\int_0^{\frac{1}{3}\pi}a^2\sin^2 3\theta d\theta$$
$$=\frac{1}{2}a^2\int_0^{\frac{1}{3}\pi}\frac{1}{2}(1-\cos 6\theta)d\theta$$
$$=\frac{1}{4}a^2\left[\theta-\frac{1}{6}\sin 6\theta\right]_0^{\frac{1}{3}\pi}$$
$$=\frac{1}{4}a^2\left(\frac{1}{3}\pi\right)=\frac{1}{12}\pi a^2$$

(ii) $y=r\sin\theta=a\sin 3\theta\sin\theta$

Max. Distance from initial line, $\frac{dy}{d\theta}=0$.

$$\frac{dy}{d\theta}=a\sin\theta(3\cos 3\theta)+a\cos\theta(\sin 3\theta)$$
$$3a\sin\theta\cos 3\theta+a\cos\theta\sin 3\theta=0$$
$$\frac{3\sin\theta\cos 3\theta}{\cos 3\theta\cos\theta}+\frac{\cos\theta\sin 3\theta}{\cos 3\theta\cos\theta}=0$$
$$3\tan\theta+\tan 3\theta=0$$

(iii)

$$3\tan\theta=\frac{3\tan\theta-tan^3\theta}{3tan^2\theta-1}$$

$$9tan^3\theta - 3\tan\theta = 3\tan\theta - tan^3\theta$$
$$10tan^3\theta - 6\tan\theta = 0$$
$$2\tan\theta\,(5tan^2\theta - 3) = 0$$
$$\tan\theta = 0\ or\ \tan\theta = \pm\sqrt{0.6}$$

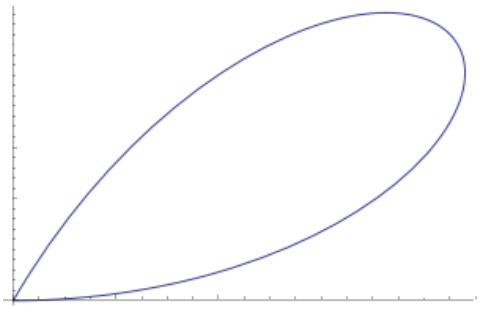
For $0 \le \theta \le \frac{1}{3}\pi$, $\tan\theta = 0$ or $\tan\theta = \frac{\sqrt{3}}{\sqrt{5}}$, hence $\sin\theta = 0\ or\ \sin\theta = \frac{\sqrt{3}}{\sqrt{8}}$.

$When\ \tan\theta = 0, \sin\theta = 0, y = 0$ (minimum distance, hence rejected).

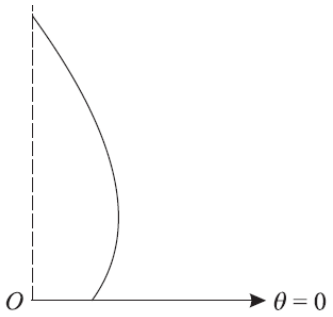
When $\tan\theta = \frac{\sqrt{3}}{\sqrt{5}}$,

$$\tan 3\theta = \frac{3\left(\sqrt{\frac{3}{5}}\right) - \sqrt{\frac{27}{125}}}{1 - \frac{9}{5}} = -\frac{3}{5}\sqrt{15}, \sin 3\theta = \sqrt{\frac{135}{160}}$$
$$y = a\sqrt{\frac{135}{160}}\sqrt{\frac{3}{8}} = \frac{9}{16}a$$

(iv)



Example 4: S2007 P1 Q5



The diagram shows the curve C with polar equation $r = e^{\theta}$, where $0 \leqslant \theta \leqslant \frac{1}{2}\pi$. Find the maximum distance of a point of C from the line $\theta = \frac{1}{2}\pi$, giving the answer in exact form. [4]

Find the area of the finite region bounded by C and the lines $\theta = 0$ and $\theta = \frac{1}{2}\pi$, giving the answer in exact form. [3]

At maximum distance, $\frac{dx}{d\theta} = 0$

$$x = rcos\theta = \cos\theta\, e^{\theta}$$
$$\frac{dx}{d\theta} = e^{\theta}\cos\theta - e^{\theta}\sin\theta$$

$\frac{dx}{d\theta} = 0, e^{\theta} \neq 0$, Hence $\cos\theta - \sin\theta = 0$.

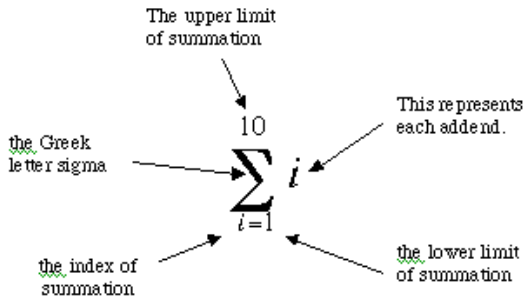
$$\tan\theta = 1, \theta = \frac{1}{4}\pi$$
$$x = \cos\frac{1}{4}\pi\, e^{\frac{1}{4}\pi} = \frac{1}{2}\sqrt{2}e^{\frac{1}{4}\pi}$$

$$Area = \frac{1}{2}\int_0^{\frac{1}{2}\pi} e^{2\theta}d\theta$$
$$= \frac{1}{2}\left[\frac{1}{2}e^{2\theta}\right]_0^{\frac{1}{2}\pi}$$
$$= \frac{1}{4}\left(e^{\pi} - 1\right)$$

CHAPTER 3: SUMMATION OF SERIES

3.1 Standard Series

A) Introducing the Summation Notation



- The Greek letter 'sigma' means sum.
- For example, the summation on the left means

$$1+2+3+4+5+6+7+8+9+10=55$$

- More examples:

$$\sum_{k=1}^{4} k^4 = 1^4+2^4+3^4+4^4 = 1+16+81+256 = 354$$

$$\sum_{m=1}^{6} \frac{1}{m} = 1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6} = \frac{49}{20}$$

$$\sum_{P=3}^{5} (P^2+1) = (3^2+1)+(4^2+1)+(5^2+1) = 53$$

$$\sum_{n=0}^{7} a_n = a_0+a_1+a_2+a_3+a_4+a_5+a_6+a_7$$

3 properties of the summation notation:

Property	Proof
$\sum_{n=1}^{k} a_n+b_n = \sum_{n=1}^{k} a_n + \sum_{n=1}^{k} b_n$	$\sum_{n=1}^{k} a_n+b_n = a_1+b_1+a_2+b_2+a_3+b_3+\ldots+a_k+b_k$ $= a_1+a_2+\ldots+a_k+b_1+b_2+\ldots+b_k$ $= \sum_{n=1}^{k} a_n + \sum_{n=1}^{k} b_n$
$\sum_{n=1}^{k} ca_n = c\sum_{n=1}^{k} a_n$, for any constant c.	$\sum_{n=1}^{k} ca_n = ca_1+ca_2+ca_3+\ldots+ca_k$ $= c(a_1+a_2+a_3+\ldots+a_k) = c\sum_{n=1}^{k} a_n$
$\sum_{n=p}^{r} a_n = \sum_{n=p}^{q} a_n + \sum_{n=q+1}^{r} a_n$, where $p<q<r$.	$\sum_{n=p}^{r} a_n = a_p+a_{p+1}+\ldots+a_{q-1}+a_q+a_{q+1}+\ldots+a_{r-1}+a_r$ $= (a_p+a_{p+1}+\ldots+a_q)+(a_{q+1}+\ldots+a_{r-1}+a_r)$ $= \sum_{n=p}^{q} a_n + \sum_{n=q+1}^{r} a_n$

B) Four Basic Results (learn all of them)

- LO: use the standard results for Σr , Σr^2 , Σr^3 to find related sums;

1st result: General	$\sum_{r=1}^{n} c = c + c + c + ... + c = cn$
2nd result (MF10):	$\sum_{r=1}^{n} r = 1 + 2 + 3 + ... + n = \frac{1}{2}n(n+1)$ Proof: Apply the arithmetic series formulae
3rd result (MF10):	$\sum_{r=1}^{n} r^2 = 1^2 + 2^2 + 3^2 + ... + n^2 = \frac{1}{6}n(n+1)(2n+1)$ Proof: shown in Section 3.2 (Telescoping series)
4th result (MF10):	$\sum_{r=1}^{n} r^3 = 1^3 + 2^3 + 3^3 + ... + n^3 = \frac{1}{4}n^2(n+1)^2 = \left(\sum_{r=1}^{n} r\right)^2$ Proof: can also be shown using telescoping series, but is shown using Mathematical Induction in Chapter 4.

Techniques for solving problems:

- Apply the 4 results above
- Break up more complex sums using the three summation rules
- Factor the expression whenever possible (don't multiply brackets unless no choice)

Example 1: W2012 P13 Q1

Show that $\sum_{r=n+1}^{2n} r^2 = \frac{1}{6}n(2n+1)(7n+1)$. [4]

$$\sum_{r=n+1}^{2n} r^2 = \sum_{r=1}^{2n} r^2 - \sum_{r=1}^{n} r^2$$

$$= \frac{1}{6}(2n)(2n+1)[2(2n)+1] - \frac{1}{6}n(n+1)(2n+1)$$

$$= \frac{1}{6}n(2n+1)[2(4n+1) - (n+1)]$$

$$= \frac{1}{6}n(2n+1)(8n+2-n-1)$$

$$= \frac{1}{6}n(2n+1)(7n+1)$$

Example 2: W2004 P1 Q5

Let

$$S_N = \sum_{n=1}^{N} (-1)^{n-1} n^3.$$

Find S_{2N} in terms of N, simplifying your answer as far as possible. [4]

Hence write down an expression for S_{2N+1} and find the limit, as $N \to \infty$, of $\frac{S_{2N+1}}{N^3}$. [3]

$$\begin{aligned}
S_{2N} &= \sum_{n=1}^{N} (-1)^{n-1} n^3 \\
&= 1^3 - 2^3 + 3^3 - 4^3 + \ldots + (2N-1)^3 - (2N)^3 \\
&= 1^3 + 2^3 + 3^3 + \ldots + (2N)^3 - 2\left[2^3 + 4^3 + \ldots + (2N)^3\right] \\
&= 1^3 + 2^3 + 3^3 + \ldots + (2N)^3 - 2^4\left(1^3 + 2^3 + \ldots + N^3\right) \\
&= \frac{1}{4}(2N)^2(2N+1)^2 - \frac{16}{4}N^2(N+1)^2 \\
&= N^2\left[(2N+1)^2 - 4(N+1)^2\right] \\
&= N^2\left(4N^2 + 4N + 1 - 4N^2 - 8N - 4\right) \\
&= -N^2(4N+3)
\end{aligned}$$

$$S_{2N+1} = -N^2(4N+3) + (2N+1)^3$$

$$\frac{S_{2N+1}}{N^3} = \left(2 + \frac{1}{N}\right)^3 - \left(4 + \frac{3}{N}\right)$$

As $N \to \infty, \frac{1}{N} \to 0$, hence $\left(2 + \frac{1}{N}\right)^3 \to 2^3$ and $\left(4 + \frac{3}{N}\right) \to 4$. Thus $S_{2N+1} \to 8 - 4 = 4$.

Example 3: W2009 P1 Q11 EITHER

Prove by induction that

$$\sum_{n=1}^{N} n^3 = \tfrac{1}{4}N^2(N+1)^2.$$ [5]

Use this result, together with the formula for $\sum_{n=1}^{N} n^2$, to show that

$$\sum_{n=1}^{N} (20n^3 + 36n^2) = N(N+1)(N+3)(5N+2).$$ [3]

Let

$$S_N = \sum_{n=1}^{N} (20n^3 + 36n^2 + \mu n).$$

Find the value of the constant μ such that S_N is of the form $N^2(N+1)(aN+b)$, where the constants a and b are to be determined. [3]

Show that, for this value of μ,

$$5 + \frac{22}{N} < N^{-4} S_N < 5 + \frac{23}{N},$$

for all $N \geqslant 18$. [3]

First part: Refer to Chapter 4: Mathematical Induction

Second part:

$$\sum_{n=1}^{N}\left(20n^3+36n^2\right)=20\sum_{n=1}^{N}n^3+36\sum_{n=1}^{N}n^2$$
$$=\frac{20}{4}N^2(N+1)^2+\frac{36}{6}N(N+1)(2N+1)$$
$$=N(N+1)\left[5N(N+1)+6(2N+1)\right]$$
$$=N(N+1)\left(5N^2+17N+6\right)=N(N+1)(N+3)(5N+2)$$

Third part:

$$S_N=\sum_{n=1}^{N}\left(20n^3+36n^2+\mu n\right)$$
$$=N(N+1)(N+3)(5N+2)+\frac{\mu}{2}N(N+1)$$
$$=N(N+1)\left(5N^2+17N+6+\frac{\mu}{2}\right)$$

When $\mu=-12$, $6+\frac{\mu}{2}=0$, hence

$$S_N=N^2(N+1)(5N+17)\Rightarrow a=5,b=17$$

Last part:

$$S_N=N^2(5N^2+22N+17)$$
$$N^{-4}S_N=5+\frac{22}{N}+\frac{17}{N^2}$$

When ≥ 18, $\frac{17}{N}<1\ \rightarrow 0<\frac{17}{N^2}<\frac{1}{N}$. Hence

$$5+\frac{22}{N}<5+\frac{22}{N}+\frac{17}{N^2}<5+\frac{22}{N}+\frac{1}{N}$$
$$5+\frac{22}{N}<N^{-4}S_N<5+\frac{23}{N}$$

3.2 Telescoping Series

- LO: use the method of differences to obtain the sum of a finite series;
- LO: recognise, by direct consideration of the sum to n terms, when a series is convergent, and find the sum to infinity in such cases.
  - The technique to sum a telescoping series is known as **method of differences**.
  - In this case, we try to express a function $g(x)$ into the form $f(x+1)-f(x)$. This allows $\sum_{x=1}^{n} g(x)$ to be evaluated as follows:

$$\begin{aligned}\sum_{x=1}^{n} g(x) &= \sum_{x=1}^{n} f(x+1)-f(x)\\ &= f(2)-f(1)+f(3)-f(2)+\ldots+f(n)-f(n-1)+f(n+1)-f(n)\\ &= f(n+1)-f(1)\end{aligned}$$

  - The colours used above indicate how equal terms are cancelled out. This method reduces the number of terms involved, allowing the series to be evaluated more easily.
  - Algebraic methods or partial fractions are usually used to change the form of the addend, but other methods may be used if appropriate.
  - Depending on the expression if $f(n+1)$ above, the summation might:

Converge, if it gets closer and closer to a certain value as n increases,
Diverge, if it approaches infinity as n gradually increases
Oscillate, if its value keeps changing in a fixed interval as *n* changes.

  - The following example is a method to prove a result stated in the previous section.

Example 1: Simplify $\left(n+\frac{1}{2}\right)^3-\left(n-\frac{1}{2}\right)^3$. Hence, show that $\sum_{r=1}^{n} r^2=\frac{1}{6}n(n+1)(2n+1)$.

$$\left(n+\frac{1}{2}\right)^3-\left(n-\frac{1}{2}\right)^3=n^3+\frac{3}{2}n^2+\frac{3}{4}n+\frac{1}{8}-\left(n^3-\frac{3}{2}n^2+\frac{3}{4}n-\frac{1}{8}\right)=3n^2+\frac{1}{4}$$

$$\begin{aligned}\sum_{r=1}^{n} 3r^2+\frac{1}{4} &= \sum_{r=1}^{n}\left(r+\frac{1}{2}\right)^3-\left(r-\frac{1}{2}\right)^3\\ &= \left(\frac{3}{2}\right)^3-\left(\frac{1}{2}\right)^3+\left(\frac{5}{2}\right)^3-\left(\frac{3}{2}\right)^3+\ldots+\left(n-\frac{1}{2}\right)^3-\left(n-\frac{3}{2}\right)^3+\left(n+\frac{1}{2}\right)^3-\left(n-\frac{1}{2}\right)^3\\ &= \left(n+\frac{1}{2}\right)^3-\frac{1}{8}\end{aligned}$$

$$\sum_{r=1}^{n} 24r^2+2=(2n+1)^3-1$$

$$24\sum_{r=1}^{n} r^2+2\sum_{r=1}^{n} 1=8n^3+12n^2+6n$$

$$24\sum_{r=1}^{n} r^2+2n=8n^3+12n^2+6n$$

$$24\sum_{r=1}^{n} r^2=4n\left(2n^2+3n+1\right)$$

$$\sum_{r=1}^{n} r^2=\frac{4}{24}n\left(2n^2+3n+1\right)$$

$$=\frac{1}{6}n(n+1)(2n+1)$$

Example 2: W2007 P1 Q2

Express

$$\frac{2n+3}{n(n+1)}$$

in partial fractions and hence use the method of differences to find

$$\sum_{n=1}^{N} \frac{2n+3}{n(n+1)}\left(\frac{1}{3}\right)^{n+1}$$

in terms of N. [4]

Deduce the value of

$$\sum_{n=1}^{\infty} \frac{2n+3}{n(n+1)}\left(\frac{1}{3}\right)^{n+1}.$$ [1]

$$\frac{2n+3}{n(n+1)} \equiv \frac{A}{n} + \frac{B}{n+1} \rightarrow 2n+3 \equiv A(n+1) + B(n)$$

$$When\ n = -1, B = -1, \qquad When\ n = 0, A = 3, \qquad Thus\ \frac{2n+3}{n(n+1)} \equiv \frac{3}{n} - \frac{1}{n+1}.$$

$$\sum_{n=1}^{N} \frac{2n+3}{n(n+1)}\left(\frac{1}{3}\right)^{n+1} = \sum_{n=1}^{N} \left(\frac{1}{3}\right)^{n+1}\left(\frac{3}{n} - \frac{1}{n+1}\right) = \sum_{n=1}^{N} \frac{1}{3^n(n)} - \frac{1}{3^{n+1}(n+1)}$$

$$= \frac{1}{3^1.1} - \frac{1}{3^2.2} + \frac{1}{3^2.2} - \frac{1}{3^3.3} + \ldots + \frac{1}{3^{N-1}(N-1)} - \frac{1}{3^N(N)} + \frac{1}{3^N(N)} - \frac{1}{3^{N+1}(N+1)}$$

$$= \frac{1}{3} - \frac{1}{3^{N+1}(N+1)}$$

As $N \rightarrow \infty, \frac{1}{3^{N+1}} \rightarrow 0$, hence $\sum_{n=1}^{\infty} \frac{2n+3}{n(n+1)}\left(\frac{1}{3}\right)^{n+1} = \frac{1}{3}$.

Example 3: W2002 P1 Q1

Given that

$$u_n = e^{nx} - e^{(n+1)x},$$

find $\sum_{n=1}^{N} u_n$ in terms of N and x. [2]

Hence determine the set of values of x for which the infinite series

$$u_1 + u_2 + u_3 + \ldots$$

is convergent and give the sum to infinity for cases where this exists. [3]

$$\sum_{n=1}^{N} u_n = \sum_{n=1}^{N} e^{nx} - e^{(n+1)x}$$

$$= e^x - e^{2x} + e^{2x} - e^{3x} + \ldots + e^{(N-1)x} - e^{Nx} + e^{Nx} - e^{(N+1)x}$$

$$= e^x - e^{(N+1)x}$$

If $x > 0$: $e^{(N+1)x} \to \infty$, hence $\sum_{n=1}^{\infty} u_n$ does not converge.

If $x < 0$: $e^{(N+1)x} \to 0$, hence $\sum_{n=1}^{\infty} u_n = e^x - 0 = e^x$.

If $x = 0$: $e^{(N+1)x} = 1$. Hence $\sum_{n=1}^{\infty} u_n = e^0 - 1 = 0$.

Example 4: S2010 P13 Q2

By considering the identity

$$\cos[(2n-1)\alpha] - \cos[(2n+1)\alpha] \equiv 2\sin\alpha\sin 2n\alpha,$$

show that if α is not an integer multiple of π then

$$\sum_{n=1}^{N} \sin(2n\alpha) = \tfrac{1}{2}\cot\alpha - \tfrac{1}{2}\operatorname{cosec}\alpha\cos[(2N+1)\alpha]. \qquad [4]$$

Deduce that the infinite series

$$\sum_{n=1}^{\infty} \sin\left(\tfrac{2}{3}n\pi\right)$$

does not converge. [1]

$$\sum_{n=1}^{N}\sin(2n\alpha)$$

$$= \frac{1}{2\sin\alpha}\sum_{n=1}^{N}\cos(2n-1)\alpha - \cos(2n+1)\alpha$$

$$= \frac{1}{2\sin\alpha}\left(\cos\alpha - \cancel{\cos 3\alpha} + \cancel{\cos 3\alpha} - \cancel{\cos 5\alpha} + \ldots\ldots + \cancel{\cos(2N-1)\alpha} - \cos(2N+1)\alpha\right)$$

$$\sum_{n=1}^{N}\sin(2n\alpha) = \frac{1}{2\sin\alpha}(\cos\alpha - \cos(2N+1)\alpha)$$

$$= \frac{1}{2}\cot\alpha - \frac{1}{2}\csc\alpha\cos(2N+1)\alpha \quad \text{(shown)}$$

Whatever the value of N, $-1 \le \cos(2N+1)\alpha \le 1$,

$$\frac{1}{2}\cot\alpha - \frac{1}{2}\csc\alpha \le \frac{1}{2}\cot\alpha - \frac{1}{2}\csc\alpha\cos(2N+1)\alpha \le \frac{1}{2}\cot\alpha + \frac{1}{2}\csc\alpha$$

Since $\sum_{n=1}^{\infty}\sin(2n\alpha)$ oscillates between $\frac{1}{2}\cot\alpha \pm \frac{1}{2}\csc\alpha$, it does not converge.

CHAPTER 4: MATHEMATICAL INDUCTION

A) The Principle of Mathematical Induction

- LO: use the method of mathematical induction to establish a given result (questions set may involve sums of series, divisibility or inequalities, for example)
- LO: recognise situations where conjecture based on a limited trial followed by inductive proof is a useful strategy, and carry this out in simple cases, e.g. to find the n^{th} derivative of xe^x .

- Use of mathematical induction: To prove that a given statement/ proposition is true for a set of integers. Often times this is a formal way to prove some statements that appear trivial.
- Two main steps are involved in a complete proof by induction:

Step 1: Base case	Verify the proposition is true for a given number (usually 1, but may vary according to situation)
Step 2: Inductive step	Prove that if the proposition is true for some positive integer k, then it is also true for $k+1$.

- Sometimes the statement is given, but we can conjecture / guess a proposition by verifying a few simple cases then proving it using mathematical induction. For example,

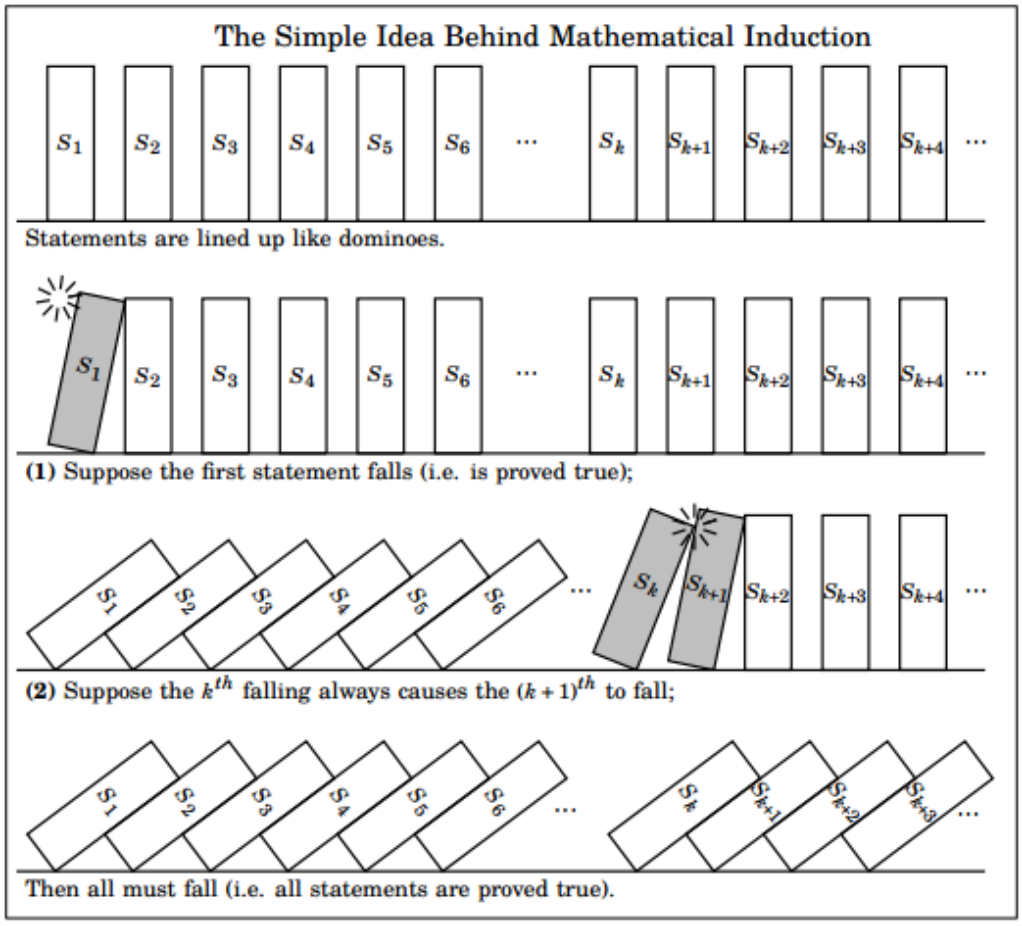
$$\frac{d}{dx}\left(xe^x\right) = xe^x + 1.e^x = (x+1)e^x$$
$$\frac{d^2}{dx^2}\left(xe^x\right) = \frac{d}{dx}\left[(x+1)e^x\right] = (x+1)e^x + 1.e^x = (x+2)e^x$$
$$\frac{d^3}{dx^3}\left(xe^x\right) = \frac{d}{dx}\left[(x+2)e^x\right] = (x+2)e^x + 1.e^x = (x+3)e^x$$

In this case, we can make a conjecture that $\frac{d^n}{dx^n}\left(xe^x\right) = (x+n)e^x$.

- The principle behind the two main steps is shown below:

Using step 1,	$P(1)$ is true.
Using step 2,	$P(2)$ is true.
Using step 2 again,	$P(3)$ is true.
Using step 2 again,	$P(4)$ is true.
…..	…
Hence if we keep using step 2,	$P(n)$ is true for all integers $n \geq 1$.

- Notice that this method only works for positive integers (for real numbers this method cannot fill in the 'gaps' between each whole number).
- Due to this property mathematical induction is always associated to dominoes, as illustrated below.


The Simple Idea Behind Mathematical Induction
Statements are lined up like dominoes.
(1) Suppose the first statement falls (i.e. is proved true);
(2) Suppose the k^{th} falling always causes the $(k+1)^{th}$ to fall;
Then all must fall (i.e. all statements are proved true).

Example 1 (Summation): We start by a formula shown in the previous chapter.

Prove by induction that $\sum_{r=1}^{n} r^3 = \frac{1}{4}n^2(n+1)^2$ for all positive integers n.

Base case:
$n = 1$ LHS: $\sum_{r=1}^{1} r^3 = 1$ RHS: $\frac{1}{4}(1)(2)^2 = 1$

LHS = RHS, hence the statement is true for $n = 1$.

Inductive step: Assume the statement is true for some integer k: $\sum_{r=1}^{k} r^3 = \frac{1}{4}k^2(k+1)^2$.

What we want to prove: when $n = k + 1$: $\sum_{r=1}^{k+1} r^3 = \frac{1}{4}(k+1)^2(k+2)^2$

$$\sum_{r=1}^{k+1} r^3 = \sum_{r=1}^{k} r^3 + (k+1)^3$$
$$= \frac{1}{4}k^2(k+1)^2 + (k+1)^3$$
$$= \frac{1}{4}(k+1)^2\left[k^2 + 4(k+1)\right] = \frac{1}{4}(k+1)^2(k+2)^2$$

Hence the statement is true for $n = k + 1$. By principle of mathematical induction,

$\sum_{r=1}^{n} r^3 = \frac{1}{4}n^2(n+1)^2$ for all positive integers n.

B) Miscellaneous Problems

- Including summation, mathematical induction can also be used to prove many different types of statements, as shown below.

Example 2: Matrices (S2011 P13 Q2)

Let $\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix}$. Prove by mathematical induction that, for every positive integer n,

$$\mathbf{A}^n = \begin{pmatrix} 2^n & 3(2^n - 1) \\ 0 & 1 \end{pmatrix}. \qquad [5]$$

Base case: $n = 1$ LHS: $\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix}$ RHS: $\begin{pmatrix} 2^1 & 3(2^1 - 1) \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix}$

LHS = RHS, hence the statement is true for $n = 1$.

Inductive step: Assume the statement is true for some integer k: $\mathbf{A}^k = \begin{pmatrix} 2^k & 3(2^k - 1) \\ 0 & 1 \end{pmatrix}$

What we want to prove: when $n = k + 1$: $\mathbf{A}^{k+1} = \begin{pmatrix} 2^{k+1} & 3(2^{k+1} - 1) \\ 0 & 1 \end{pmatrix}$

$$\begin{aligned} \mathbf{A}^{k+1} &= \mathbf{A}^k \mathbf{A} \\ &= \begin{pmatrix} 2^k & 3(2^k - 1) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2(2^k) + 0\left(3(2^k - 1)\right) & 3(2^k) + 3(2^k - 1) \\ 0(2) + 1(0) & 0(3) + 1(1) \end{pmatrix} \\ &= \begin{pmatrix} 2(2^k) & 3(2.2^k - 1) \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2^{k+1} & 3(2^{k+1} - 1) \\ 0 & 1 \end{pmatrix} \end{aligned}$$

Hence the statement is true for $n = k + 1$. By principle of mathematical induction, $\mathbf{A}^n = \begin{pmatrix} 2^n & 3(2^n - 1) \\ 0 & 1 \end{pmatrix}$ for all positive integers n.

Example 3: Divisibility (S2006 P1 Q3)

- General strategy for divisibility problems: consider $f(n + 1) - kf(n)$, where k is any integer that is not a multiple of the number you want to prove it is divisible by.
- For example, if we want to prove that $f(n)$ is divisible by 11, we can try $f(n + 1) + f(n)$ or $f(n + 1) - 3f(n)$, but trying $f(n + 1) - 11f(n)$ would not help.

Q: Prove by induction that $23^{2n} + 31^{2n} + 46$ is divisible by 48, for all integers $n \geq 0$. [6]

Let $f(n) = 23^{2n} + 31^{2n} + 46$.

Base case: $n = 0$ $f(0) = 23^0 + 31^0 + 46 = 48$.

Hence the statement is true for $n = 0$.

Inductive step:

$$\begin{aligned} f(n+1) &= 23^{2(n+1)} + 31^{2(n+1)} + 46 \\ &= 529(23^{2n}) + 961(31^{2n}) + 46 \\ f(n+1) - f(n) &= 528\left(23^{2n}\right) + 960\left(31^{2n}\right) \\ f(n+1) &= f(n) + 48\left[20\left(31^{2n}\right) + 11\left(23^{2n}\right)\right] \end{aligned}$$

Since $f(n+1)-f(n)$ is divisible by 48, if $f(n)$ is divisible by 48, so does $f(n+1)$. Hence by principle of mathematical induction, $f(n)$ is divisible by 48 for all positive integers n.

Example 4: Inequalities (S2007 P1 Q8)

- General strategy for inequalities: to prove $x_{n+1} > k$, prove $x_{n+1} - k > 0$.

The sequence $x_1, x_2, x_3, \ldots$ is such that $x_1 = 1$ and

$$x_{n+1} = \frac{1+4x_n}{5+2x_n}.$$

Prove by induction that $x_n > \frac{1}{2}$ for all $n \geqslant 1$. [5]

Prove also that $x_n > x_{n+1}$ for all $n \geqslant 1$. [3]

Base case: $n = 1$ $x_1 = 1 > \frac{1}{2}$ (given) , hence the statement is true for $n = 1$.

Inductive step: Assume the statement is true for some integer k: $x_k > \frac{1}{2}$

What we want to prove: when $n = k+1$: $x_{k+1} > \frac{1}{2}$

$$x_{k+1} = \frac{1+4x_k}{5+2x_k}$$

$$x_{k+1} - \frac{1}{2} = \frac{1+4x_k}{5+2x_k} - \frac{1}{2}$$

$$= \frac{2(1+4x_k) - (5+2x_k)}{2(5+2x_k)}$$

$$= \frac{6x_k - 3}{2(5+2x_k)} = \frac{3\left(x_k - \frac{1}{2}\right)}{5+2x_k}$$

If $x_k > \frac{1}{2}$, then $x_k - \frac{1}{2} > 0$ and $5 + 2x_k > 0$, hence $x_{k+1} - \frac{1}{2} > 0$, $x_{k+1} > \frac{1}{2}$.

Hence the statement is true for $n = k+1$. By principle of mathematical induction, $x_n > \frac{1}{2}$ for all positive integers n.

$$x_n - x_{n+1} = x_n - \frac{1+4x_n}{5+2x_n}$$

$$= \frac{5x_n + 2x_n^2 - 1 - 4x_n}{5+2x_n}$$

$$= \frac{2x_n^2 + x_n - 1}{5+2x_n} = \frac{(2x_n - 1)(x_n + 1)}{5+2x_n}$$

As $x_n > \frac{1}{2}, 2x_n - 1 > 0, x_n + 1 > 0, 5 + 2x_n > 0$.

Hence $x_n - x_{n+1} > 0$, $x_n > x_{n+1}$.

Example 5: Differentiation (S2004 P1 Q4)

➢ This one starts by forming the proposition to prove using a conjecture.

It is given that

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{\ln x}{x}\right) = \frac{a_n \ln x + b_n}{x^{n+1}},$$

where a_n and b_n depend only on n.

(i) Find a_1, a_2 and a_3. [3]

(ii) Use mathematical induction to establish a formula for a_n. [5]

(i)

$$\frac{d}{dx}\left(\frac{\ln x}{x}\right) = \frac{x\left(\frac{1}{x}\right) - 1(\ln x)}{x^2} = \frac{-\ln x + 1}{x^2}$$

$$\frac{d^2}{dx^2}\left(\frac{\ln x}{x}\right) = \frac{d}{dx}\left(\frac{-\ln x + 1}{x^2}\right) = \frac{x^2\left(-\frac{1}{x}\right) - (1-\ln x)(2x)}{x^4} = \frac{2\ln x - 3}{x^3}$$

$$\frac{d^3}{dx^3}\left(\frac{\ln x}{x}\right) = \frac{d}{dx}\left(\frac{2\ln x - 3}{x^3}\right) = \frac{x^3\left(\frac{2}{x}\right) - (2\ln x - 3)\left(3x^2\right)}{x^6} = \frac{-6\ln x + 11}{x^4}$$

Comparing coefficients, we have $a_1 = -1, a_2 = 2, a_3 = -6$.

(ii) For safety purposes, we will verify another case before conjecturing:

$$\frac{d^4}{dx^4}\left(\frac{\ln x}{x}\right) = \frac{d}{dx}\left(\frac{-6\ln x + 11}{x^4}\right) = \frac{x^4\left(-\frac{6}{x}\right) - (-6\ln x + 11)\left(4x^3\right)}{x^8} = \frac{24\ln x - 50}{x^5}$$

Comparing coefficients, $a_4 = 24$, thus we conjecture that $a_n = (-1)^n n!$.

Base case: $n = 1$ (Verified true in the previous part)

Inductive step: Assume the statement is true for some integer k, $a_k = (-1)^k k!$.

To prove: for $n = k + 1$: $a_{n+1} = (-1)^{n+1}(n+1)!$.

$$\frac{d^{k+1}}{dx^{k+1}}\left(\frac{\ln x}{x}\right) = \frac{d}{dx}\left(\frac{a_k \ln x + b_k}{x^{k+1}}\right)$$

$$= \frac{x^{k+1}\left(\frac{a_k}{x}\right) - (k+1)(a_k \ln x + b_k)\left(x^k\right)}{x^{2k+2}}$$

$$= \frac{a_k - (k+1)(a_k \ln x + b_k)}{x^{k+2}}$$

$$= \frac{-(k+1)a_k \ln x + \left[a_k - (k+1)b_k\right]}{x^{k+2}}$$

Comparing coefficients,

$$a_{k+1} = -(k+1)a_k$$
$$= -(k+1)(-1)^n k!$$
$$= (-1)^{n+1}(k+1)!$$

Hence by principle of mathematical induction, $a_n = (-1)^n n!$ for all positive integers n.

Example 6: Inequalities (W2004 P1 Q8)

- For inequalities of these type 'bounds' need to introduced bravely. Sometimes working backwards helps one to see the conditions to be inserted.

The sequence of real numbers $a_1, a_2, a_3, \ldots$ is such that $a_1 = 1$ and

$$a_{n+1} = \left(a_n + \frac{1}{a_n}\right)^{\lambda},$$

where λ is a constant greater than 1. Prove by mathematical induction that, for $n \geqslant 2$,

$$a_n \geqslant 2^{g(n)},$$

where $g(n) = \lambda^{n-1}$. [6]

Prove also that, for $n \geqslant 2$, $\frac{a_{n+1}}{a_n} > 2^{(\lambda-1)g(n)}$. [3]

Base case: $n = 2$

$$a_2 = \left(a_1 + \frac{1}{a_1}\right)^{\lambda} = 2^{\lambda} = 2^{g(2)},$$

hence the statement is true for $n = 2$.

Inductive step: Assume the statement is true for some integer k: $a_k \geq 2^{g(k)}$

What we want to prove: when $n = k + 1$: $a_{k+1} \geq 2^{g(k+1)}$ $\left[a_{k+1} \geq 2^{(\lambda^k)}\right]$

$$\begin{aligned} a_{k+1} &= \left(a_k + \frac{1}{a_k}\right)^{\lambda} \\ &> (a_k + 0)^{\lambda} \\ &\geq \left[2^{g(k)}\right]^{\lambda} \\ &= \left[2^{\left(\lambda^{k-1}\right)}\right]^{\lambda} = 2^{\lambda^k} \end{aligned}$$

Since $a_{k+1} \geq 2^{g(k+1)}$, the statement is true for $n = k + 1$. By principle of mathematical induction, $a_n \geq 2^{g(n)}$ for all positive integers n.

From above we know that $a_{n+1} > a_n^{\lambda}$, hence

$$\frac{a_{n+1}}{a_n} > a_n^{\lambda-1} \geq 2^{\lambda-1 g(n)}$$

Example 7: Recursive Formulae

- Sometimes the number of base cases we need to verify will vary, for example here we need to verify 2 as this is a second order recurrence relation.

A sequence of numbers $u_1, u_2, u_3, \ldots$ is such that $u_1 = u_2 = 1$ and $u_{n+2} = u_{n+1} + u_n$ for $n \geq 1$. Prove that for all $n \geq 1$,

$$u_n = \frac{(1+\sqrt{5})^n - (1-\sqrt{5})^n}{2^n\sqrt{5}}.$$

Base case: $n = 1$

$$u_1 = \frac{(1+\sqrt{5}) - (1-\sqrt{5})}{2\sqrt{5}} = \frac{2\sqrt{5}}{2\sqrt{5}} = 1.$$

Base case: $n = 2$

$$\left(1+\sqrt{5}\right)^2 = 1 + 2\sqrt{5} + 5 = 6 + 2\sqrt{5}$$
$$\left(1-\sqrt{5}\right)^2 = 1 - 2\sqrt{5} + 5 = 6 - 2\sqrt{5}$$

$$u_2 = \frac{\left(1+\sqrt{5}\right)^2 - \left(1-\sqrt{5}\right)^2}{2^2\sqrt{5}}$$
$$= \frac{6 + 2\sqrt{5} - 6 + 2\sqrt{5}}{4\sqrt{5}}$$
$$= \frac{4\sqrt{5}}{4\sqrt{5}} = 1$$

Hence the statement is true for $n = 1$ and $n = 2$.

Inductive step: Assume the statement is true for some integers k and $k + 1$:

$$u_k = \frac{(1+\sqrt{5})^k - (1-\sqrt{5})^k}{2^k\sqrt{5}}, \quad u_{k+1} = \frac{(1+\sqrt{5})^{k+1} - (1-\sqrt{5})^{k+1}}{2^{k+1}\sqrt{5}}$$

What we want to prove: when $n = k + 2$: $u_{k+2} = \frac{(1+\sqrt{5})^{k+2} - (1-\sqrt{5})^{k+2}}{2^{k+2}\sqrt{5}}$.

$$u_{k+2} = u_{k+1} + u_k$$
$$= \frac{\left(1+\sqrt{5}\right)^{k+1} - \left(1-\sqrt{5}\right)^{k+1}}{2^{k+1}\sqrt{5}} + \frac{\left(1+\sqrt{5}\right)^{k} - \left(1-\sqrt{5}\right)^{k}}{2^{k}\sqrt{5}}$$
$$= \frac{\left(1+\sqrt{5}\right)^{k+1} - \left(1-\sqrt{5}\right)^{k+1} + 2\left(1+\sqrt{5}\right)^{k} - 2\left(1-\sqrt{5}\right)^{k}}{2^{k+1}\sqrt{5}}$$
$$= \frac{\left(1+\sqrt{5}\right)^{k}\left(1+\sqrt{5}+2\right) - \left(1-\sqrt{5}\right)^{k}\left(1-\sqrt{5}+2\right)}{2^{k+1}\sqrt{5}}$$
$$= \frac{\frac{1}{2}\left(1+\sqrt{5}\right)^{k}\left(6+2\sqrt{5}\right) - \frac{1}{2}\left(1-\sqrt{5}\right)^{k}\left(6-2\sqrt{5}\right)}{2^{k+1}\sqrt{5}}$$
$$= \frac{\left(1+\sqrt{5}\right)^{k}\left(1+\sqrt{5}\right)^2 - \left(1-\sqrt{5}\right)^{k}\left(1-\sqrt{5}\right)^2}{2^{k+2}\sqrt{5}} = \frac{\left(1+\sqrt{5}\right)^{k+2} - \left(1-\sqrt{5}\right)^{k+2}}{2^{k+2}\sqrt{5}}$$

Hence the statement is true for $n = k + 2$. By principle of mathematical induction, $u_n = \frac{(1+\sqrt{5})^n - (1-\sqrt{5})^n}{2^n\sqrt{5}}$ for all positive integers n.

CHAPTER 5: DIFFERENTIATION AND INTEGRATION

5.1 New Differentiation Techniques

- LO: obtain an expression for $\frac{d^2y}{dx^2}$ in cases where the relation between *y* and *x* is defined implicitly or parametrically;

A) Pre-Requisite Knowledge: Differentiation (Paper 1 and Paper 3)

- Derivatives of basic functions: Polynomial, exponents, logarithms, trigonometry
- Product rule, quotient rule and chain rule
- Applications: increasing / decreasing functions, maximum and minimum points, normal and tangent equations, rate of change etc.

B) Implicit Differentiation

- In Mathematics Paper 3 you only use this to find $\frac{dy}{dx}$.
- In FM you will extend this technique to find $\frac{d^2y}{dx^2}$ and maybe higher derivatives.
- For expressions in terms of *x* only, differentiate normally.
- For expressions in terms of *y* only, apply the chain rule.
- For expressions in terms of both *x* and *y*, apply the product rule or quotient rule.

Example: A curve has equation $x^3 + y^3 = x\ln y$. Find $\frac{dy}{dx}$ and $\frac{d^2y}{dx^2}$.

Differentiating with respect to *x*:

$$3x^2 + 3y^2\frac{dy}{dx} = \frac{x}{y}\frac{dy}{dx} + \ln y$$

Rearranging: $$3x^2y + 3y^3\tfrac{dy}{dx} = x\tfrac{dy}{dx} + y\ln y - (1)$$

$$\frac{dy}{dx} = \frac{y\ln y - 3x^2y}{3y^3 - x}$$

Method 1: Quotient rule

$$\frac{d^2y}{dx^2} = \frac{(3y^3 - x)\frac{d}{dx}(y\ln y - 3x^2y) - (y\ln y - 3x^2y)\frac{d}{dx}(3y^3 - x)}{(3y^3 - x)^2}$$

$$= \frac{(3y^3 - x)\left(\ln y\ \frac{dy}{dx} + \frac{dy}{dx} - 6xy - 3x^2\frac{dy}{dx}\right) - (y\ln y - 3x^2y)\left(9y^2\frac{dy}{dx} - 1\right)}{(3y^3 - x)^2}$$

Lastly regroup terms and substitute $\frac{dy}{dx}$ in if needed.

$$\frac{d^2y}{dx^2}(3y^3 - x)^2 = \frac{dy}{dx}(3y^3 - x)(\ln y\ + 1 - 3x^2) - 6xy - (y\ln y - 3x^2y)\left(9y^2\frac{dy}{dx}\right) + (y\ln y - 3x^2y)$$

$$= -6xy - \frac{9y^2(y\ln y - 3x^2y)^2}{3y^3 - x} + \frac{(y\ln y - 3x^2y)[(3y^3 - x) + (3y^3 - x)(\ln y\ + 1 - 3x^2)]}{3y^3 - x}$$

$$-6xy - \frac{9y^2(y\ln y - 3x^2y)^2}{3y^3 - x} + (y\ln y - 3x^2y)[1 + (\ln y\ + 1 - 3x^2)]$$

$$\frac{d^2y}{dx^2} = \frac{-9y^2(y\ln y - 3x^2y)^2}{(3y^3 - x)^3} + \frac{(\ln y + 2 - 3xy)(y\ln y - 3x^2y)}{(3y^3 - x)^2} - \frac{6xy}{3y^3 - x}$$

Method 2: Differentiate (1) with respect to x again:

$$3x^2y + 3y^3\frac{dy}{dx} = x\frac{dy}{dx} + y\ln y$$

$$6xy + 3x^2\frac{dy}{dx} + 9y^2\left(\frac{dy}{dx}\right)^2 + 3y^3\frac{d^2y}{dx^2} = x\frac{d^2y}{dx^2} + \frac{dy}{dx} + \ln y\frac{dy}{dx} + \frac{dy}{dx}$$

$$6xy + 3x^2\frac{dy}{dx} + 9y^2\left(\frac{dy}{dx}\right)^2 + 3y^3\frac{d^2y}{dx^2} = x\frac{d^2y}{dx^2} + 2\frac{dy}{dx} + \ln y\frac{dy}{dx}$$

Lastly regroup terms and substitute $\frac{dy}{dx}$ in if needed.

$$\frac{d^2y}{dx^2}(3y^3 - x) = -9y^2\left(\frac{dy}{dx}\right)^2 + \frac{dy}{dx}(\ln y + 2 - 3xy) - 6xy$$

$$= -9y^2\frac{(y\ln y - 3x^2y)^2}{(3y^3 - x)^2} + \frac{(\ln y + 2 - 3xy)(y\ln y - 3x^2y)}{3y^3 - x} - 6xy$$

$$\frac{d^2y}{dx^2} = \frac{-9y^2(y\ln y - 3x^2y)^2}{(3y^3 - x)^3} + \frac{(\ln y + 2 - 3xy)(y\ln y - 3x^2y)}{(3y^3 - x)^2} - \frac{6xy}{3y^3 - x}$$

- Method 2 is preferred as less simplification is required.
- The example appears messy, will be greatly simplified when numbers are plugged in, as seen below.

Example 1: W2002 P1 Q6 (Uses Method 1)

A curve has equation $x^3 + xy^2 - y^3 = 3$.

(i) Show that there is no point of the curve at which $\frac{dy}{dx} = 0$. [4]

(ii) Find the values of $\frac{dy}{dx}$ and $\frac{d^2y}{dx^2}$ at the point $(1, -1)$. [5]

(i) Differentiating with respect to x:

$$3x^2 + y^2 + x\frac{d}{dx}(y^2) - \frac{d}{dx}(y^3) = 0$$

$$3x^2 + y^2 + 2xy\frac{dy}{dx} - 3y^2\frac{dy}{dx} = 0$$

$$\frac{dy}{dx} = \frac{3x^2 + y^2}{3y^2 - 2xy}$$

When $\frac{dy}{dx} = 0$, $3x^2 + y^2 = 0$. Then only values of (x, y) that satisfies it are $(0,0)$.
Substituting into the curve, we get $0 = 3$, hence $(0,0)$ does not lie on the curve.
Hence there are no points on the curve that satisfies $\frac{dy}{dx} = 0$.

(ii) At $(1, -1)$:

$$\frac{dy}{dx} = \frac{3+1}{3+2} = \frac{4}{5}$$

$$\frac{d^2y}{dx^2} = \frac{(3y^2 - 2xy)\frac{d}{dx}(3x^2 + y^2) - (3x^2 + y^2)\frac{d}{dx}(3y^2 - 2xy)}{(3y^2 - 2xy)^2}$$

$$= \frac{(3y^2 - 2xy)\left(6x + 2y\frac{dy}{dx}\right) - (3x^2 + y^2)\left(6y\frac{dy}{dx} - 2x\frac{dy}{dx} - 2y\right)}{(3y^2 - 2xy)^2}$$

Substituting $x = 1, y = -1$ and $\frac{dy}{dx} = \frac{4}{5}$:

$$\frac{d^2y}{dx^2} = \frac{(3+2)\left[6 - 2\left(\frac{4}{5}\right)\right] + 4\left[-8\left(\frac{4}{5}\right) + 2\right]}{(3+2)^2} = \frac{5(30-8) - 4(10-32)}{5^3} = \frac{198}{125}$$

Example 2: W2006 P1 Q10 (Uses Method 2)

The curve C has equation

$$y = x^2 + \lambda \sin(x + y),$$

where λ is a constant, and passes through the point $A\left(\frac{1}{4}\pi, \frac{1}{4}\pi\right)$. Show that C has no tangent which is parallel to the y-axis. [5]

Show that, at A,

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = 2 - \tfrac{1}{64}\pi(4-\pi)(\pi+2)^2. \quad [5]$$

$$y = x^2 + \lambda \sin(x + y)$$

Differentiating with respect to x:

$$\frac{dy}{dx} = 2x + \lambda \cos(x+y)\frac{d}{dx}(x+y)$$

$$\frac{dy}{dx} = 2x + \lambda \cos(x+y)\left(1 + \frac{dy}{dx}\right) \ (*)$$

Parallel to the y-axis implies $\frac{dy}{dx} = \infty$. Divide the whole equation by $\frac{dy}{dx}$:

$$1 = \lambda \cos(x+y)$$

Curve passes through $\left(\frac{1}{4}\pi, \frac{1}{4}\pi\right)$, hence $1 = \lambda \cos\left(\frac{\pi}{2}\right), 1 = 0$. This is a contradiction, hence C cannot have a tangent which is parallel to the y-axis.

At $\left(\frac{1}{4}\pi, \frac{1}{4}\pi\right)$, $\frac{1}{4}\pi = \frac{1}{16}\pi^2 + \lambda$ and $\frac{dy}{dx} = \frac{1}{2}\pi$.

Differentiating (*) with respect to x:

$$\frac{d^2y}{dx^2} = 2 + \lambda \frac{d}{dx}\left[\cos(x+y)\left(1 + \frac{dy}{dx}\right)\right]$$

$$\frac{d^2y}{dx^2} = 2 + \lambda \cos(x+y)\frac{d}{dx}\left(1 + \frac{dy}{dx}\right) + \left(1 + \frac{dy}{dx}\right)\lambda \frac{d}{dx}\cos(x+y)$$

$$\frac{d^2y}{dx^2} = 2 + \lambda \cos(x+y)\frac{d^2y}{dx^2} - \sin(x+y)\left(1 + \frac{dy}{dx}\right)\lambda \frac{d}{dx}(x+y)$$

$$\frac{d^2y}{dx^2} = 2 + \lambda \cos(x+y)\frac{d^2y}{dx^2} - \lambda \sin(x+y)\left(1 + \frac{dy}{dx}\right)^2$$

When $x = \frac{1}{4}\pi, y = \frac{1}{4}\pi, \frac{dy}{dx} = \frac{1}{2}\pi$:

$$\frac{d^2y}{dx^2} = 2 - \lambda \sin\left(\frac{\pi}{2}\right)\left(1 + \frac{\pi}{2}\right)^2 = 2 - \frac{1}{4}\lambda(2+\pi)^2 = 2 - \left(\frac{1}{16}\pi - \frac{1}{64}\pi^2\right)(2+\pi)^2$$

$$\frac{d^2y}{dx^2} = 2 - \frac{1}{64}\pi(4-\pi)(2+\pi)^2$$

C) Parametric Differentiation

- In Mathematics Paper 3 you only use this to find $\frac{dy}{dx}$.
- In FM you will extend this technique to find $\frac{d^2y}{dx^2}$ and maybe higher derivatives
- Pre-requisite knowledge from P3: If $x = x(t)$ and $y = y(t)$, by chain rule:

$$\boldsymbol{\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{y'(t)}{x'(t)}}$$

- To obtain $\frac{d^2y}{dx^2}$, we let $z = \frac{dy}{dx}$, hence $\frac{dz}{dx} = \frac{d^2y}{dx^2}$, and apply the rule above by changing y to z:

$$\boldsymbol{\frac{dz}{dx} = \frac{\frac{dz}{dt}}{\frac{dx}{dt}} = \frac{\frac{d}{dt}\left(\frac{dy}{dx}\right)}{\frac{dx}{dt}} = \frac{y''(t)x'(t) - x''(t)y'(t)}{[x'(t)]^3}}$$

Example 1: S2002 P1 Q7

Given that

$$x = 1 + \frac{1}{t} \quad \text{and} \quad y = t^3 e^{-t},$$

where $t \neq 0$, find $\frac{dy}{dx}$ in terms of t. [3]

Find $\frac{d^2y}{dx^2}$ in terms of t, and hence find the values of t for which $\frac{d^2y}{dx^2} = 0$. [5]

$$x = 1 + \frac{1}{t} \rightarrow \frac{dx}{dt} = -\frac{1}{t^2}$$

$$y = t^3 e^{-t} \rightarrow y = e^{-t}(3t^2 - t^3)$$

$$\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = t^2 e^{-t}(3t^2 - t^3) = t^4 e^{-t}(3 - t)$$

$$\frac{d^2y}{dx^2} = -t^2 \frac{d}{dt}\left[e^{-t}\left(3t^4 - t^5\right)\right]$$

$$= -t^2 e^{-t}\left(12t^3 - 5t^4 - 3t^4 + t^5\right) = t^{-3} e^{-t}\left(8t - t^2 - 12\right)$$

$$\frac{d^2y}{dx^2} = 0, e^{-t} \neq 0, t \neq 0, \text{hence}$$

$$t^2 - 8t + 12 = 0$$

$$(t - 6)(t - 2) = 0$$

$$t = 6 \; or \; t = 2$$

Example 2: W2009 P1 Q4

It is given that

$$x = t + \sin t, \quad y = t^2 + 2\cos t,$$

where $-\pi < t < \pi$. Find $\frac{dy}{dx}$ in terms of t. [2]

Show that

$$\frac{d^2y}{dx^2} = \frac{2t\sin t}{(1+\cos t)^3}.$$ [4]

Show that $\frac{dy}{dx}$ increases with x over the given interval of t. [2]

$$x = t + \sin t \rightarrow \frac{dx}{dt} = 1 + \cos t$$

$$y = t^2 + 2\cos t \rightarrow \frac{dy}{dt} = 2t - 2\sin t$$

$$\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{2t - 2\sin t}{1 + \cos t}$$

$$\begin{aligned}
\frac{d^2y}{dx^2} &= \frac{\frac{d}{dt}\left(\frac{2t-2\sin t}{1+\cos t}\right)}{1+\cos t} \\
&= \frac{(1+\cos t)(2-2\cos t) - (2t-2\sin t)(-\sin t)}{(1+\cos t)^2(1+\cos t)} \\
&= \frac{2(1+\cos t)(1-\cos t) + 2\sin t(t-\sin t)}{(1+\cos t)^3} \\
&= \frac{2(1-\cos^2 t) + 2t\sin t - 2\sin^2 t}{(1+\cos t)^3} \\
&= \frac{2t\sin t}{(1+\cos t)^3}
\end{aligned}$$

In $-\pi < t < \pi$, $1 + \cos t > 0$, hence $(1 + \cos t)^3 > 0$.

In $0 < t < \pi$, $t > 0, \sin t > 0$, hence $2t \sin t > 0$.

In $-\pi < t < 0$, $t > 0, \sin t > 0$, hence $2t \sin t > 0$.

Since $2t \sin t > 0$ and $(1 + \cos t)^3 > 0$, $\frac{2t \sin t}{(1+\cos t)^3} > 0$.

Since $\frac{d^2y}{dx^2} > 0$ for all values of t in the interval,

$\frac{dy}{dx}$ increases over x over the given interval of t (except when $t = 0$).

5.2 New Integration Techniques

- LO: derive and use reduction formulae for the evaluation of definite integrals in simple cases;

A) Review: Old Integration Techniques

- Integrals: Polynomials, Exponents, Trigonometry, Fractions
- Indefinite and Definite integrals, and their properties
- Applications of Integration: Area under a curve, Volume of revolution
- Integration by Recognition / Using identities:

1. $\int \tan^2 x dx = \int \sec^2 x - 1 dx = \tan x - x + c$
2. $\int \sin^2 x dx = \int \frac{1}{2}(1 - \cos 2x) dx = \frac{1}{2}x - \frac{1}{4}\sin 2x + c$
3. $\int \cos^2 x dx = \int \frac{1}{2}(1 + \cos 2x) dx = \frac{1}{2}x + \frac{1}{4}\sin 2x + c$
4. $\int \tan x dx = -\int \frac{-\sin x}{\cos x} dx = -\ln|\cos x| + c$ OR

   $\int \tan x dx = \int \frac{\sec x \tan x}{\sec x} dx = \ln|\sec x| + c$

   (Note that both expressions are equal)
5. $\int \sec x dx = \int \frac{\sec^2 x + \sec x \tan x}{\sec x + \tan x} dx = \ln|\sec x + \tan x| + c$

- Integration by Substitution (replace both the integrand and the limits):

1. Using $x = \tan u$, find $\int \frac{1}{1+x^2} dx$

$$x = \tan u \rightarrow dx = sec^2 u du$$

$$\int \frac{1}{1+x^2} dx = \int \frac{sec^2 u du}{1 + tan^2 u} = \int du = u + c = \tan^{-1} x + c$$

2. Using $u = \sin x$, find $\int_0^{\frac{1}{2}\pi} \sin^2 2x \cos x dx$

$$u = \sin x \rightarrow du = \cos x \, dx$$

$$x = 0, u = \sin 0 = 0, \qquad x = \frac{1}{2}\pi, u = \sin \frac{1}{2}\pi = 1$$

$$\int_0^{\frac{1}{2}\pi} sin^2 2x \cos x \, dx = \int_0^1 sin^2 2x \, du = \int_0^1 4 sin^2 x cos^2 x \, du$$

$$= 4\int_0^1 u^2(1-u^2) \, du = 4\int_0^1 u^2 - u^4 \, du$$

$$= 4\left[\frac{1}{3}u^3 - \frac{1}{5}u^5\right]_0^1 = 4\left(\frac{1}{3} - \frac{1}{5}\right) = \frac{8}{15}$$

3. Using $u = ax + b$,

$$u = ax + b \rightarrow du = a\,dx$$

(i) $$\int \frac{1}{ax+b}\,dx = \int \frac{1}{u}\left(\frac{dx}{a}\right) = \frac{1}{a}\ln|u| + c = \frac{1}{a}\ln|ax+b| + c$$

(ii) $$\int \csc^2(ax+b)dx = \int \csc^2 u\left(\frac{du}{a}\right) = \frac{1}{a}\cot u + c = \frac{1}{a}\cot(ax+b) + c$$

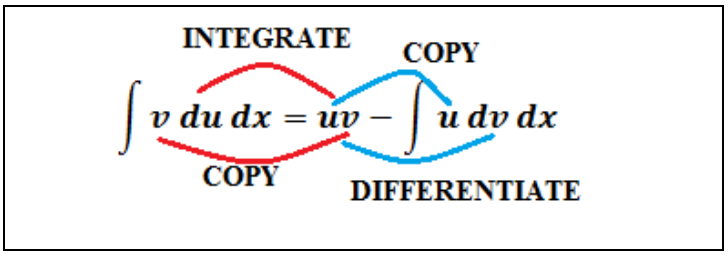
4. By using the substitution $u = f(x)$, show that

(i) $$\int \frac{f'(x)}{f(x)}\,dx = \ln|f(x)| + c$$

(ii) $$\int f'(x)[f(x)]^n\,dx = \frac{[f(x)]^{n+1}}{n+1} + c$$

(iii) $$\int f'(x)e^{f(x)}dx = e^{f(x)} + c$$

- Integration by Parts:



- Integration by parts is used when the integrand is a product of two functions.
- It is essentially the integration version of the product rule.
- To utilise it, we choose a function to be integrated and a function to be differentiated.
- The acronym above is what I used to avoid listing '*v*' and '*du*' to simplify my work done, but be careful that the two times 'COPY' stated above means two different actions!
- Generally functions involving exponents, logarithms and polynomials are chosen to be differentiated, but it might vary from situation to situation.
- Choosing the correct function to differentiate or integrate will simplify the problem, while choosing the wrong one will make the problem tougher, so practice a lot as experience is the best method to determine which function must be differentiated or integrated.

1. $$\int_0^{\frac{1}{2}\pi} x\sin x dx = \left[-x\cos x\right]_0^{\frac{1}{2}\pi} + \int_0^{\frac{1}{2}\pi} \cos x.1 dx$$
$$= \int_0^{\frac{1}{2}\pi} \cos x dx \quad .$$
$$= \left[\sin x\right]_0^{\frac{1}{2}\pi} = 1$$
In this case: we choose x to differentiate and $\sin x$ to integrate.

2. $$\int \ln x dx = x\ln x - \int x\left(\frac{1}{x}\right)dx = x\ln x - x + c$$
In this case: we choose $\ln x$ to differentiate and 1 to integrate.

3. This problem is unique: the choice of the function to be integrated or differentiated does not matter.
In this case: we need to perform it twice, as we cannot eliminate either of the functions by differentiating them.
$$\int e^x \cos x dx = e^x \cos x + \int e^x \sin x dx$$
$$= e^x \cos x + e^x \sin x - \int e^x \cos x dx$$
$$2\int e^x \cos x dx = e^x(\cos x + \sin x) + 2c \Rightarrow \int e^x \cos x dx = \frac{1}{2}e^x(\cos x + \sin x) + c$$
Note: The constant added at the back can vary (I added $2c$ to allow the final answer be in terms of c)

4. This is a more complicated one, as we are splitting it up using identities.
$$\int \tan^3 x dx = \int \tan x\left(\tan^2 x\right)dx$$
$$= \int \tan x\left(\sec^2 x - 1\right)dx$$
$$= \int \tan x \sec^2 x dx - \int \tan x dx$$
Then we evaluate the first integral by parts (differentiating $\tan x$ and integrating $\sec^2 x$), or use the substitution $u = \tan x$:
$$\int \tan x \sec^2 x dx = \tan x.\tan x - \int \tan x \sec^2 x$$
$$2\int \tan x \sec^2 x dx = \tan^2 x$$
Then substitute back to solve the integral.
$$\int \tan^3 x dx = \int \tan x \sec^2 x dx - \int \tan x dx$$
$$= \frac{1}{2}\tan^2 x + \int \frac{-\sin x}{\cos x} dx = \frac{1}{2}\tan^2 x + \ln|\cos x| + c$$

B) Reduction Formula

- This is the new integration technique for FM
- Involves reducing an integral into a simpler one using a recursion formula
- Questions mainly involve deriving and using them to find integrals.
- Main strategies involved are integration by parts (ensure that you are confident with them) and reverse differentiation (will be shown in the examples). Sometimes integration by substitution will be used too.

Example 1: Find a reduction formula for $sin^n\theta$. Hence evaluate $\int sin^4\theta\ d\theta$ and $\int sin^5\theta\ d\theta$.

Let $I_n = \int sin^n\theta\ d\theta$. Split the integral and perform integration by parts (we choose to integrate $\sin\theta$ and differentiate $sin^{n-1}\theta$) :

$$I_n = \int \sin\theta \sin^{n-1}\theta d\theta$$
$$= -\cos\theta(\sin^{n-1}\theta) + \int \cos\theta(n-1)\sin^{n-2}\theta\cos\theta d\theta$$

Then use $sin^2\theta + cos^2\theta = 1$:

$$I_n = -\cos\theta(\sin^{n-1}\theta) + (n-1)\int \cos^2\theta \sin^{n-2}\theta d\theta$$
$$= -\cos\theta(\sin^{n-1}\theta) + (n-1)\int \left(1-\sin^2\theta\right)\sin^{n-2}\theta d\theta$$
$$= -\cos\theta(\sin^{n-1}\theta) + (n-1)\int \sin^{n-2}\theta - \sin^n\theta d\theta$$

This allows us to rearrange for the final result:

$$I_n = -\cos\theta(\sin^{n-1}\theta) + (n-1)\left(I_{n-2} - I_n\right)$$
$$nI_n = (n-1)I_{n-2} - \cos\theta(\sin^{n-1}\theta)$$

To find I_4, we first find I_0 since it's simpler:

$$I_0 = \int sin^0\theta\ d\theta = \int d\theta = \theta + c$$

Then use the reduction formula. Let $n = 2$ to find I_2:

$$2I_2 = I_0 - \cos\theta \sin\theta \rightarrow I_2 = \frac{1}{2}\theta - \frac{1}{2}\cos\theta \sin\theta + c$$

Then let $n = 4$ to find I_4. Notice that we do not move the constant over as adding one in the end is more convenient.

$$4I_4 = 3I_2 - \cos\theta\ sin^3\theta \rightarrow \int sin^4\theta\ d\theta = \frac{3}{8}\theta - \frac{3}{8}\cos\theta \sin\theta - \frac{1}{4}\cos\theta\ sin^3\theta + c$$

To find I_5, instead of I_0 we have to find I_1.

$$I_1 = \int \sin\theta\ d\theta = -\cos\theta + c$$

Let $n = 3$:

$$3I_3 = 2I_1 - \cos\theta\ sin^2\theta \rightarrow I_3 = -\frac{2}{3}\cos\theta - \frac{1}{3}\cos\theta\ sin^2\theta + c$$

Then let $n = 5$.

$$5I_5 = 4I_3 - \cos\theta\ sin^4\theta$$
$$\int sin^5\theta\ d\theta = -\frac{8}{15}\cos\theta - \frac{4}{5}\cos\theta\ sin^2\theta - \frac{1}{5}\cos\theta\ sin^4\theta + c$$

Comments: This is a 2-step reduction formula. Notice how we need different starting values to find the integrals depending on whether *n* is odd or even.

Example 2: Let $I_n = \int_1^e ln^n x\ dx$. Show that $I_n = (n-1)(I_{n-2} - I_{n-1}), n \geq 2$ and evaluate $\int_1^e ln^3 x\ dx$.[Note: $ln^n x = (\ln x)^n$]

First we need to evaluate $\int \ln x\, dx$:

$$\int \ln x\, dx = x\ln x - \int x\left(\frac{1}{x}\right)dx = x\ln x - x + c$$

Then we split the integrand and apply integration by parts.

$$\int_1^e \ln^n x\,dx = \int_1^e (\ln x)\ln^{n-1} x\,dx$$
$$= \left[\ln^{n-1} x(x\ln x - x)\right]_1^e - \int_1^e (x\ln x - x)(n-1)\ln^{n-2} x\left(\frac{1}{x}\right)dx$$
$$I_n = -(n-1)\int_1^e (\ln x - 1)\ln^{n-2} x\,dx$$
$$= (n-1)\int_1^e \ln^{n-2} x - \ln^{n-1} x\,dx$$
$$I_n = (n-1)(I_{n-2} - I_{n-1})$$

This is a double reduction formula. Hence to start we need both I_0 and I_1:

$$I_0 = \int_1^e ln^0 x\,dx = e - 1 \qquad I_1 = \int_1^e \ln x\,dx = e - e + 1 = 1$$

Hence let $n = 2$: $I_2 = I_0 - I_1 = e - 2$

$n = 3$: $I_3 = 2I_1 - 2I_2 = 4 - 2e$

Example 3: W2011 P11 Q6 (this one is a one-step reduction)

Let $I_n = \int_0^1 x^n(1-x)^{\frac{1}{2}}\,dx$, for $n \geqslant 0$. Show that, for $n \geqslant 1$,

$$(3 + 2n)I_n = 2nI_{n-1}. \qquad [5]$$

Hence find the exact value of I_3. [3]

$$I_n = \int_0^1 x^n(1-x)^{\frac{1}{2}}\,dx = \left[\frac{(1-x)^{\frac{3}{2}}}{-\frac{3}{2}}(x^n)\right]_0^1 + \frac{2}{3}\int_0^1 nx^{n-1}(1-x)^{\frac{3}{2}}\,dx$$
$$3I_n = 2n\int_0^1 x^{n-1}(1-x)^{\frac{1}{2}}(1-x)dx$$
$$3I_n = 2n\int_0^1 x^{n-1}(1-x)^{\frac{1}{2}}dx - 2n\int_0^1 x^n(1-x)^{\frac{1}{2}}dx$$
$$3I_n = 2nI_{n-1} - 2nI_n \rightarrow (3+2n)I_n = 2nI_{n-1}$$
$$I_0 = \int_0^1 (1-x)^{\frac{1}{2}}\,dx = \left[\frac{(1-x)^{\frac{3}{2}}}{-\frac{3}{2}}\right]_0^1 = \frac{2}{3}$$

When $n = 1$: $5I_1 = 2I_0 \rightarrow I_1 = \frac{4}{15}$

When $n = 2$: $7I_2 = 4I_1 \rightarrow I_2 = \frac{16}{105}$

When $n = 3$: $9I_3 = 6I_2 \rightarrow I_3 = \frac{32}{315}$

Example 4: S2002 P1 Q2 (this one interchanges between two reduction formulae)

Given that

$$C_n = \int_0^1 (1-x)^n \cos x\,dx \quad \text{and} \quad S_n = \int_0^1 (1-x)^n \sin x\,dx,$$

show that, for $n \geqslant 1$,

$$C_n = nS_{n-1} \quad \text{and} \quad S_n = 1 - nC_{n-1}. \qquad [3]$$

Hence find the value of S_3, correct to 6 decimal places. [3]

$$C_n = \int_0^1 (1-x)^n \cos x\,dx$$
$$= \left[(1-x)^n \sin x\right]_0^1 + \int_0^1 n(1-x)^{n-1} \sin x\,dx$$
$$= n\int_0^1 (1-x)^{n-1} \sin x\,dx = nS_{n-1}$$

$$S_n = \int_0^1 (1-x)^n \sin x\,dx$$
$$= \left[(1-x)^n(-\cos x)\right]_0^1 - \int_0^1 n(1-x)^{n-1}\cos x\,dx$$
$$= [0-1(-\cos 0)] - n\int_0^1 (1-x)^{n-1}\cos x\,dx$$
$$= 1 - nC_{n-1}$$

To find S_3 we have to interchange between the two formulae.

$n = 3$: $S_3 = 1-3C_2 \rightarrow C_2 = \frac{1}{3}(1-S_3)$

$n = 2$: $C_2 = 2S_1 \rightarrow S_1 = \frac{1}{6}(1-S_3)$

$n = 1$: $S_1 = 1-C_0$, $C_0 = 1-\frac{1}{6}(1-S_3)$

$$S_3 = 6C_0 - 5$$
$$= 6\int_0^1 \cos x\,dx - 5$$
$$= 6\left[\sin x\right]_0^1 - 5$$
$$= 6(\sin 1) - 5 = 0.048826$$

Example 5: W2008 P1 Q7 (this one uses the reverse-differentiation technique)

Let $I_n = \int_0^1 \frac{1}{(1+x^4)^n}\,dx$. By considering $\frac{d}{dx}\left(\frac{x}{(1+x^4)^n}\right)$, show that

$$4nI_{n+1} = \frac{1}{2^n} + (4n-1)I_n. \qquad [4]$$

Given that $I_1 = 0.866\,97$, correct to 5 decimal places, find I_3. [4]

$$\frac{d}{dx}\left[\frac{x}{(1+x^4)^n}\right] = \frac{(1+x^4)^n - x(n)(1+x^4)^{n-1}(4x^3)}{(1+x^4)^{2n}}$$
$$= \frac{1+x^4-4nx^4}{(1+x^4)^{n+1}}$$
$$= \frac{1+(1-4n)(x^4+1)-1+4n}{(1+x^4)^{n+1}}$$
$$= \frac{4n}{(1+x^4)^{n+1}} + \frac{1-4n}{(1+x^4)^n}$$

Integrating both sides over the interval [0,1]:

$$\left[\frac{x}{(1+x^4)^n}\right]_0^1 = 4n\int_0^1 \frac{1}{(1+x^4)^{n+1}}dx + (1-4n)\int_0^1 \frac{1}{(1+x^4)^n}dx$$

$$\frac{1}{2^n} = 4nI_{n+1} + (1-4n)I_n \rightarrow 4nI_{n+1} = \frac{1}{2^n} + (4n-1)I_n$$

When $n = 1$:

$$4I_2 = \frac{1}{2} + 3I_1 \rightarrow I_2 = \frac{1}{8} + \frac{3}{4}I_1$$

When $n = 2$:

$$8I_3 = \frac{1}{4} + 7I_2$$

$$I_3 = \frac{1}{32} + \frac{7}{8}\left(\frac{1}{8} + \frac{3}{4}I_1\right)$$

$$= \frac{9}{64} + \frac{21}{32}(0.86697) \approx 0.70957$$

Example 6: S2010 P11 Q5 (proving reduction formulae using mathematical induction)

Let

$$I_n = \int_1^e x(\ln x)^n \, dx,$$

where $n \geqslant 1$. Show that

$$I_{n+1} = \tfrac{1}{2}e^2 - \tfrac{1}{2}(n+1)I_n. \qquad [3]$$

Hence prove by induction that, for all positive integers n, I_n is of the form $A_n e^2 + B_n$, where A_n and B_n are rational numbers. [6]

$$I_{n+1} = \int_1^e x(\ln x)^{n+1} dx$$

$$= \left[(\ln x)^{n+1}\left(\frac{1}{2}x^2\right)\right]_1^e - \int_1^e \left(\frac{1}{2}x^2\right)(n+1)(\ln x)^n \left(\frac{1}{x}\right) dx$$

$$= \frac{1}{2}e^2 - \frac{1}{2}(n+1)\int_1^e x(\ln x)^n \, dx$$

$$= \frac{1}{2}e^2 - \frac{1}{2}n(n+1)I_n$$

Base case: $n = 1$

$$I_1 = \int_1^e x \ln x \, dx = \left[\left(\frac{1}{2}x^2\right)\ln x\right]_1^e - \int_1^e \left(\frac{1}{2}x^2\right)\left(\frac{1}{x}\right) dx$$

$$= \frac{1}{2}e^2 - \int_1^e \frac{1}{2}x \, dx$$

$$= \frac{1}{4}e^2 + \frac{1}{4}$$

Since $A_1 = B_1 = \frac{1}{4}$, the statement is true for $n = 1$.

Inductive step: Assume the statement is true for some integer k, hence $I_k = A_k e^2 + B_k$.

To prove: for $n = k + 1$: $I_{k+1} = A_{k+1}e^2 + B_{k+1}$

$$I_{k+1} = \frac{1}{2}e^2 - \frac{1}{2}(k+1)I_k$$

$$= \frac{1}{2}e^2 - \frac{1}{2}(k+1)\left(A_k e^2 + B_k\right)$$

$$= e^2\left[\frac{1}{2} - \frac{1}{2}(k+1)A_k\right] - \frac{1}{2}(k+1)B_k$$

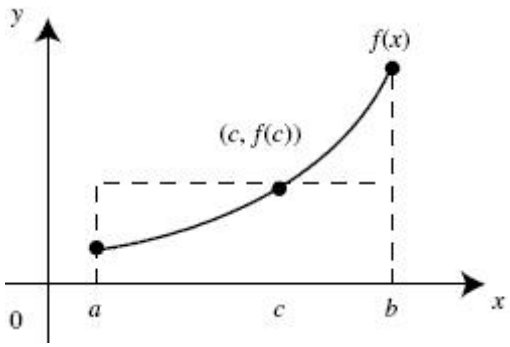
Since A_k and B_k are rational numbers, so is $\frac{1}{2} - \frac{1}{2}(k+1)A_k$ and $-\frac{1}{2}(k+1)B_k$. Hence A_{k+1} and B_{k+1} are also rational numbers. The statement is true for $n = k + 1$. Thus by principle of mathematical induction, $I_n = A_n e^2 + B_n$ for all integers n.

5.3 Applications of Integration

- LO: use integration to find:
  - mean values and centroids of two- and three dimensional figures (where equations are expressed in cartesian coordinates, including the use of a parameter), using strips, discs or shells as appropriate,
  - arc lengths (for curves with equations in cartesian coordinates, including the use of a parameter, or in polar coordinates),
  - surface areas of revolution about one of the axes (for curves with equations in cartesian coordinates, including the use of a parameter, but not for curves with equations in polar coordinates).

  - In Maths Paper 1 and Paper 3 you used integration to find area and volume of revolution. Now you will be introduced to more applications of it.

A) Mean Value



- The mean value theorem: If $f(x)$ is continuous on $[a, b]$, there exists an number c in $[a, b]$ such that

$$\boxed{\int_a^b \boldsymbol{f(x)\,dx} = \boldsymbol{f(c)(b-a)} \rightarrow \boldsymbol{f(c)} = \frac{\mathbf{1}}{\boldsymbol{b-a}}\int_a^b \boldsymbol{f(x)\,dx}}$$

- $f(c)$ is known as the mean value of the function.
- Main concept involved: think of the area under the curve as a rectangle. Thus

$$\text{Area} = \text{Width} \times \text{Mean Value} \rightarrow \int_a^b f(x)\,dx = f(c)(b-a)$$

Example: W2008 P1 Q2

Let $y = e^x$. Find the mean value of y with respect to x over the interval $0 \leqslant x \leqslant 2$. [2]

Show that the mean value of x with respect to y over the interval $1 \leqslant y \leqslant e^2$ is $\frac{e^2+1}{e^2-1}$. [4]

$$\text{Area} = \int_0^2 e^x\,dx = e^2 - 1$$
$$\text{Width} = 2 - 0 = 2$$
$$y_m(\text{Width}) = \text{Area}$$
$$y_m = \frac{1}{2}(e^2 - 1)$$

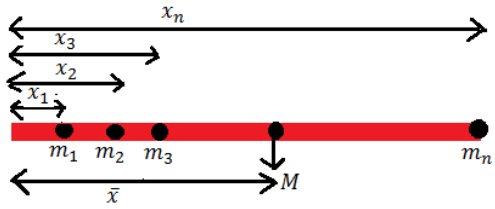
$$y = e^x \rightarrow x = \ln y$$
$$Area = \int_1^{e^2} \ln y\,dy$$
$$= \left[y\ln y - y\right]_1^{e^2}$$
$$= 2e^2 - e^2 + 1$$
$$= e^2 + 1$$

$$\text{Width} = e^2 - 1$$
$$x_m(\text{Width}) = \text{Area}$$
$$x_m = \frac{e^2+1}{e^2-1}$$

B) Centroid / Centre of Mass

- A centroid or centre of mass is where the weight of a lamina / solid is assumed to be acted on.
- Before starting this section, you should have some basic knowledge about moments and mass, either from Topic 12 in the Mechanics section or from AS-Level Physics.

For one-dimensional objects: Volume does not exist, thus: Mass = Density × Length



Think of the rod as *n* point masses.

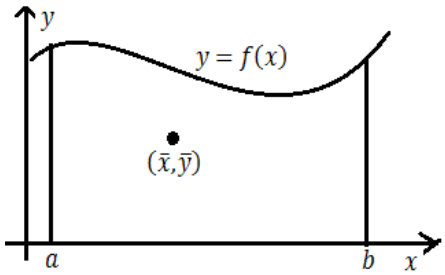
Taking moments from one end of the rod:

$$\bar{x} = \frac{m_1x_1 + m_2x_2 + \cdots + m_nx_n}{M}$$

For two-dimensional objects:

Assumption: Density of the material is the same throughout the lamina.
For a two dimensional object, volume does not exist, thus: Mass = Density × Area
Below let M_x and M_y be the 'moment' equations for the system along the x and y axis

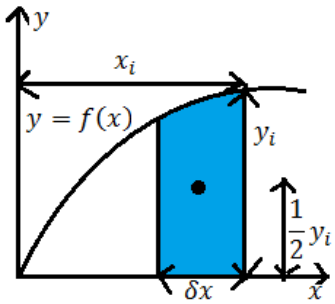
$$\left(M_x = \frac{\text{Moment}}{\text{Density}}\right)$$



Taking moments along the *y*-axis:

$$\rho M_x = M\bar{x}$$
$$\rho M_x = \rho A\bar{x}$$
$$\bar{x} = \frac{\rho M_x}{\rho A} = \frac{M_x}{A}$$

Similarly, $\bar{y} = \frac{M_y}{A}$.



Think of the area as n small rectangular strips of fixed widths δx and variable lengths y_i.

Taking moments about the y-axis:

M_x = Sum of all Areas × Distance

$$= \sum y_i \delta x(x_i)$$

Note: The distance from x-axis does not matter as δx is very small.

When δx is small, $M_x \to \int xy\,dx$

$$\boldsymbol{\bar{x} = \frac{M_x}{\text{Area}} = \frac{\int xy\,dx}{\int y\,dx}}$$

Taking moments about the y-axis:

M_x = Sum of all Areas × Distance

$$= \sum y_i \delta x\left(\frac{1}{2}y_i\right)$$

Note: The distance from y-axis MATTERS.

When δx is small, $M_y \to \int \frac{1}{2}y^2\,dx$

$$\boldsymbol{\bar{y} = \frac{M_y}{\text{Area}} = \frac{\int \frac{1}{2}y^2\,dx}{\int y\,dx}}$$

Example: W2005 P1 Q8

Find the coordinates of the centroid of the finite region bounded by the x-axis and the curve whose equation is

$$y = x^2(1-x). \qquad [7]$$

Deduce the coordinates of the centroid of the finite region bounded by the x-axis and the curve whose equation is

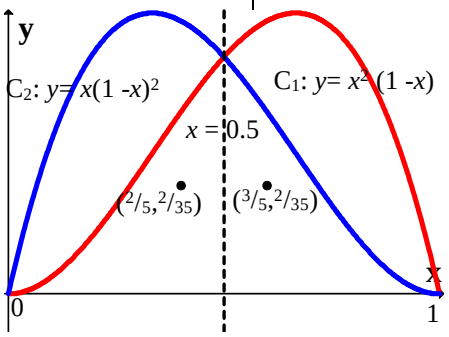
$$y = x(1-x)^2. \qquad [2]$$

$$M_x = \int_0^1 x^3(1-x)dx = \left[\frac{1}{4}x^4 - \frac{1}{5}x^5\right]_0^1 = \frac{1}{20}$$

$$M_y = \frac{1}{2}\int_0^1 x^4(1-x)^2dx = \frac{1}{2}\left[\frac{1}{5}x^5 - \frac{2}{6}x^6 + \frac{1}{7}x^7\right]_0^1 = \frac{1}{210}$$

$$\text{Area} = \int_0^1 x^2(1-x)dx = \left[\frac{1}{3}x^3 - \frac{1}{4}x^4\right]_0^1 = \frac{1}{12}$$

$$(\bar{x}, \bar{y}) = \left(\frac{12}{20}, \frac{12}{210}\right) = \left(\frac{3}{5}, \frac{2}{35}\right)$$



C_2 is obtained when C_1 is reflected along the line $x = 0.5$.

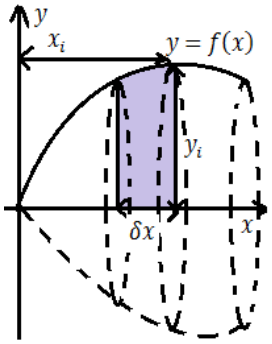
Applying the transformation $f(x) = 1 - x$ for the new centroid:

$$\left(\frac{3}{5}, \frac{2}{35}\right) \to \left(1 - \frac{3}{5}, \frac{2}{35}\right) = \left(\frac{2}{5}, \frac{2}{35}\right)$$

For three-dimensional objects:

- Recall that the volume of a cylinder is $\pi r^2 h$.

x-coordinate:



By symmetry, $\bar{y} = 0$.

Think of each strip as a cylinder with height δx and radius y_i.

Taking moments about the *y*-axis:

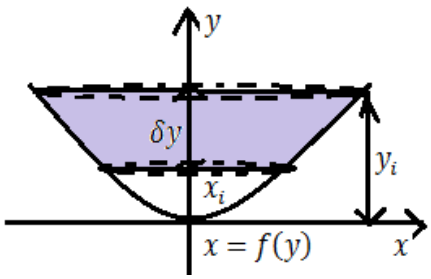
M_x = Sum of all volumes × Distance

$$= \sum \pi y_i^2 \delta x(x_i)$$

When δx is small, $M_x \to \int \pi x y^2\, dx$

$$\bar{x} = \frac{\boldsymbol{M_x}}{\textbf{Volume}} = \frac{\int \boldsymbol{\pi x y^2\, dx}}{\int \boldsymbol{\pi y^2\, dx}} = \frac{\int \boldsymbol{x y^2\, dx}}{\int \boldsymbol{y^2\, dx}}$$

y-coordinate:



By symmetry, $\bar{x} = 0$.

Think of each strip as a cylinder with height δy and radius x_i.

Taking moments about the *x*-axis:

M_y = Sum of all volumes × Distance

$$= \sum \pi x_i^2 \delta y(y_i)$$

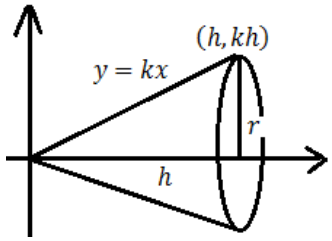
When δx is small, $M_y \to \int \pi x^2 y\, dy$

$$\bar{y} = \frac{\boldsymbol{M_y}}{\textbf{Volume}} = \frac{\int \boldsymbol{\pi x^2 y\, dy}}{\int \boldsymbol{\pi x^2\, dy}} = \frac{\int \boldsymbol{x^2 y\, dy}}{\int \boldsymbol{x^2\, dy}}$$

Example: S2008 P1 Q1

The finite region enclosed by the line $y = kx$, where k is a positive constant, the x-axis for $0 \leqslant x \leqslant h$, and the line $x = h$ is rotated through 1 complete revolution about the x-axis. Prove by integration that the centroid of the resulting cone is at a distance $\frac{3}{4}h$ from the origin O. [4]

[The volume of a cone of height h and base radius r is $\frac{1}{3}\pi r^2 h$.]

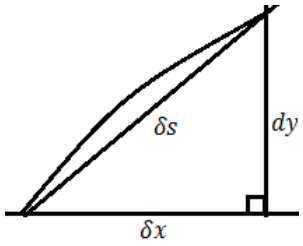


By symmetry, $\bar{y} = 0$.

$$M_x = \int_0^h \pi x y^2 dx$$
$$= \pi k^2 \int_0^h x^3 dx$$
$$= \frac{1}{4}\pi k^2 h^4$$

$$r = kh \rightarrow k = \frac{r}{h}$$
$$M_y = \frac{1}{4}\pi k^2 h^4 = \frac{1}{4}\pi r^2 h^2$$
$$\bar{y} = \frac{\mathrm{M_y}}{\text{Volume}} = \frac{\frac{1}{4}\pi r^2 h^2}{\frac{1}{3}\pi r^2 h} = \frac{3}{4}h$$
$$d = \sqrt{(\bar{x} - 0)^2 + (\bar{y} - 0)^2} = \frac{3}{4}h$$

C) Arc Length



- You will need to know all 3 cases: Cartesian, parametric and polar curves.
- Learn the diagram on the left.
- Using Pythagoras' theorem:

$$(\delta s)^2 = (\delta x)^2 + (\delta y)^2$$

Case 1: When $y = f(x)$

$$\left(\frac{\delta s}{\delta x}\right)^2 = 1 + \left(\frac{\delta y}{\delta x}\right)^2$$

When $\delta x \rightarrow 0$,

$$\frac{\delta s}{\delta x} \rightarrow \frac{ds}{dx} \text{ and } \frac{\delta y}{\delta x} \rightarrow \frac{dy}{dx}:$$
$$\left(\frac{ds}{dx}\right)^2 = 1 + \left(\frac{dy}{dx}\right)^2$$
$$\frac{ds}{dx} = \sqrt{1 + \left(\frac{dy}{dx}\right)^2}$$
$$\boldsymbol{s = \int \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx}$$

Case 2: When $x = f(y)$

$$\left(\frac{\delta s}{\delta y}\right)^2 = 1 + \left(\frac{\delta x}{\delta y}\right)^2$$

When $\delta y \rightarrow 0$,

$$\frac{\delta s}{\delta y} \rightarrow \frac{ds}{dy} \text{ and } \frac{\delta x}{\delta y} \rightarrow \frac{dx}{dy}:$$
$$\left(\frac{ds}{dy}\right)^2 = 1 + \left(\frac{dx}{dy}\right)^2$$
$$\frac{ds}{dy} = \sqrt{1 + \left(\frac{dx}{dy}\right)^2}$$
$$\boldsymbol{s = \int \sqrt{1 + \left(\frac{dx}{dy}\right)^2}\, dy}$$

Case 3: Parametric equations

$$x = x(t), \quad y = y(t)$$
$$\left(\frac{\delta s}{\delta t}\right)^2 = \left(\frac{\delta x}{\delta t}\right)^2 + \left(\frac{\delta y}{\delta t}\right)^2$$

When $\delta t \rightarrow 0$,

$$\frac{\delta s}{\delta t} \rightarrow \frac{ds}{dt}, \frac{\delta x}{\delta t} \rightarrow \frac{dx}{dt}, \frac{\delta y}{\delta t} \rightarrow \frac{dy}{dt}$$

$$\left(\frac{ds}{dt}\right)^2 = \left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2$$
$$\frac{ds}{dt} = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}$$
$$\boldsymbol{s = \int \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt}$$

Case 4: Polar coordinates

$$x = r\cos\theta \rightarrow \frac{dx}{d\theta} = \frac{dr}{d\theta}\cos\theta - r\sin\theta \quad -(1)$$

$$y = r\sin\theta \rightarrow \frac{dy}{d\theta} = \frac{dr}{d\theta}\sin\theta + r\cos\theta \quad -(2)$$

$$\left(\frac{dx}{d\theta}\right)^2 + \left(\frac{dy}{d\theta}\right)^2 = \left(\frac{dr}{d\theta}\right)^2 \cos^2\theta - r\frac{dr}{d\theta}\cos\theta\sin\theta + r^2\sin^2\theta$$
$$+ \left(\frac{dr}{d\theta}\right)^2 \sin^2\theta + r\frac{dr}{d\theta}\cos\theta\sin\theta + r^2\cos^2\theta$$
$$= \left(\frac{dr}{d\theta}\right)^2 \left(\cos^2\theta + \sin^2\theta\right) + r^2\left(\cos^2\theta + \sin^2\theta\right)$$
$$= r^2 + \left(\frac{dr}{d\theta}\right)^2$$

Substitute into Case 3, replacing the parameter from t to θ:

$$s = \int \sqrt{\left(\frac{dx}{d\theta}\right)^2 + \left(\frac{dy}{d\theta}\right)^2}\, d\theta$$

$$\boxed{s = \int \sqrt{r^2 + \left(\frac{dr}{d\theta}\right)^2}\, d\theta}$$

Example 1: W2009 P1 Q8(a)

The curve C_1 has equation $y = -\ln(\cos x)$. Show that the length of the arc of C_1 from the point where $x = 0$ to the point where $x = \frac{1}{3}\pi$ is $\ln(2+\sqrt{3})$. [5]

$$y = -\ln\cos x$$
$$\frac{dy}{dx} = -\frac{(-\sin x)}{\cos x} = \tan x$$
$$1 + \left(\frac{dy}{dx}\right)^2 = 1 + tan^2x = sec^2x$$

$$s = \int_0^{\frac{1}{3}\pi} \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx$$
$$= \int_0^{\frac{1}{3}\pi} \sec x\, dx$$
$$= \left[\ln|\sec x + \tan x|\right]_0^{\frac{1}{3}\pi}$$
$$= \ln|2+\sqrt{3}| - \ln 1 = \ln(2+\sqrt{3})$$

Example 2: W2004 P1 Q4

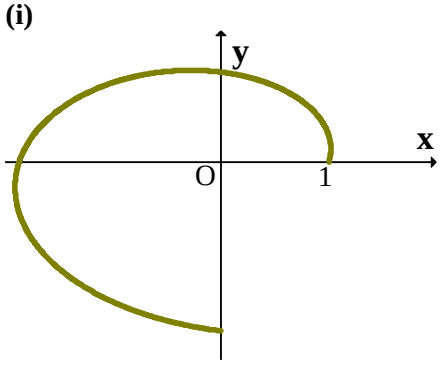
The curve C has polar equation

$$r = e^{\frac{1}{5}\theta}, \qquad 0 \leqslant \theta \leqslant \tfrac{3}{2}\pi.$$

(i) Draw a sketch of C. [2]

(ii) Find the length of C, correct to 3 significant figures. [4]

(i)



(ii)

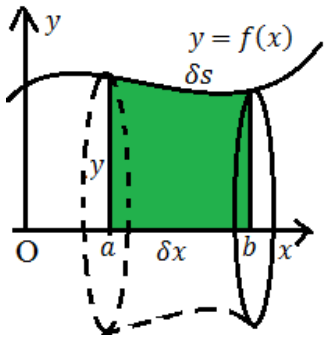
$$r^2 + \left(\frac{dr}{d\theta}\right)^2 = e^{\frac{2}{5}\theta} + \left(\frac{1}{5}e^{\frac{1}{5}\theta}\right)^2 = \frac{26}{25}e^{\frac{2}{5}\theta}$$
$$s = \int_0^{\frac{3}{2}\pi} \sqrt{r^2 + \left(\frac{dr}{d\theta}\right)^2}\, d\theta$$
$$= \frac{\sqrt{26}}{5}\int_0^{\frac{3}{2}\pi} e^{\frac{1}{5}\theta}\, d\theta$$
$$= \sqrt{26}\left[e^{0.2\theta}\right]_0^{\frac{3}{2}\pi}$$
$$= \sqrt{26}\left(e^{0.3\pi} - 1\right) \approx 7.99$$

D) Surface Area of Revolution

- Recall from the previous section that

$$\frac{ds}{dx} = \sqrt{1 + \left(\frac{dy}{dx}\right)^2} - (1), \frac{ds}{dy} = \sqrt{1 + \left(\frac{dx}{dy}\right)^2} - (2), \quad \frac{ds}{dt} = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} - (3)$$

For rotation about the *x*-axis:



Think of each strip as a cylinder with height δs and radius y.

Circumference, $C = 2\pi y$.

Surface Area = Cirumference × Height

$$= \sum C\delta s = \sum 2\pi y \delta s$$

When δs is small, Surface Area, S:

$$\boldsymbol{S = \int 2\pi y\, ds}$$

Using (1), we get the Cartesian form:

$$\boldsymbol{S = \int 2\pi y \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx}$$

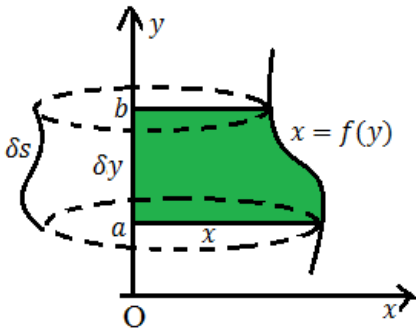
Using (2):

$$\boldsymbol{S = \int 2\pi y \sqrt{1 + \left(\frac{dx}{dy}\right)^2}\, dy}$$

Using (3), we get the Parametric Form:

$$\boldsymbol{S = \int 2\pi y \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt}$$

For rotation about the *y*-axis:



Think of each strip as a cylinder with height δs and radius x.

Circumference, $C = 2\pi x$.

Surface Area = Cirumference × Height

$$= \sum C\delta s = \sum 2\pi x \delta s$$

When δs is small, Surface Area, S:

$$\boldsymbol{S = \int 2\pi x\, ds}$$

Using (1), we get the Cartesian form:

$$\boldsymbol{S = \int 2\pi x \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx}$$

Using (2):

$$\boldsymbol{S = \int 2\pi x \sqrt{1 + \left(\frac{dx}{dy}\right)^2}\, dy}$$

Using (3), we get the Parametric Form:

$$\boldsymbol{S = \int 2\pi x \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt}$$

Not in syllabus: Let $x = r\cos\theta$, $y = r\sin\theta$ and use the 4th form of ds to obtain the equation for surface area in polar form:

$$\boldsymbol{S = \int 2\pi r\, cos\theta \sqrt{r^2 + \left(\frac{dr}{d\theta}\right)^2}\, d\theta}, \qquad \boldsymbol{S = \int 2\pi r\, sin\theta \sqrt{r^2 + \left(\frac{dr}{d\theta}\right)^2}\, d\theta},$$

Example 1: W2009 P1 Q8(b)

The curve C_2 has equation $y = 2\sqrt{(x+3)}$. The arc of C_2 joining the point where $x = 0$ to the point where $x = 1$ is rotated through one complete revolution about the x-axis. Show that the area of the surface generated is

$$\tfrac{8}{3}\pi(5\sqrt{5}-8). \qquad [5]$$

$$y = 2(x+3)^{\frac{1}{2}}$$
$$\frac{dy}{dx} = (x+3)^{-\frac{1}{2}}$$
$$1+\left(\frac{dy}{dx}\right)^2 = 1+\frac{1}{x+3} = \frac{x+4}{x+3}$$

$$S = \int_0^1 2\pi y\sqrt{1+\left(\frac{dy}{dx}\right)^2}\,dx$$
$$= 2\pi\int_0^1 2\sqrt{x+3}\sqrt{\frac{x+4}{x+3}}dx$$
$$= 4\pi\int_0^1 \sqrt{x+4}dx$$
$$= 4\pi\left[\frac{2}{3}(x+4)^{\frac{3}{2}}\right]_0^1$$
$$= \frac{8}{3}\pi\left(5^{1.5}-4^{1.5}\right)$$
$$= \frac{8}{3}\pi\left(5\sqrt{5}-8\right)$$

Example 2: W2012 P13 Q6

The curve C has parametric equations

$$x = t^2, \quad y = \tfrac{1}{4}t^4 - \ln t,$$

for $1 \leqslant t \leqslant 2$. Find the area of the surface generated when C is rotated through 2π radians about the y-axis. [7]

$$x = t^2 \rightarrow \frac{dx}{dt} = 2t$$
$$y = \frac{1}{4}t^4 - \ln t \rightarrow \frac{dy}{dt} = t^3 - \frac{1}{t}$$
$$\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 = 4t^2 + \left(t^3 - \frac{1}{t}\right)^2$$
$$= \left(t^3 + \frac{1}{t}\right)^2$$

$$S = \int_1^2 2\pi x\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\,dt$$
$$= 2\pi\int_1^2 t^2\left(t^3 + \frac{1}{t}\right)dt$$
$$= 2\pi\int_1^2 t^5 + t\,dt$$
$$= 2\pi\left[\frac{1}{6}t^6 + \frac{1}{2}t^2\right]_1^2$$
$$= 2\pi\left(\frac{64}{6} + \frac{4}{2} - \frac{1}{6} - \frac{1}{2}\right)$$
$$= 24\pi$$

5.4 Miscellaneous Problems

- Typical exam questions that combine several parts of this chapter are shown here.

Example 1: S2003 P1 Q7

The variables x and y are related by the equation $x^4 + y^4 = 1$, where $0 < x < 1$ and $0 < y < 1$.

(i) Obtain an equation which relates x, y, $\frac{dy}{dx}$, $\frac{d^2y}{dx^2}$, and deduce that

$$\frac{d^2y}{dx^2} = -\frac{3x^2}{y^7}.$$ [6]

(ii) Given that $y = b_1$ when $x = a_1$ and that $y = b_2$ when $x = a_2$, where $a_1 < a_2$, prove that the mean value of $\frac{d^3y}{dx^3}$ with respect to x over the interval $a_1 \leqslant x \leqslant a_2$ is

$$\frac{3(a_1^2 b_2^7 - a_2^2 b_1^7)}{b_1^7 b_2^7 (a_2 - a_1)}.$$ [4]

(i)

$$x^4 + y^4 = 1$$

$$4x^3 + 4y^3 \frac{dy}{dx} = 0$$

$$x^3 + y^3 \frac{dy}{dx} = 0 \rightarrow \frac{dy}{dx} = -\frac{x^3}{y^3}$$

$$3x^2 + 3y^2 \left(\frac{dy}{dx}\right)^2 + y^3 \frac{d^2y}{dx^2} = 0$$

$$3x^2 + 3\frac{x^6}{y^4} + y^3 \frac{d^2y}{dx^2} = 0$$

$$y^3 \frac{d^2y}{dx^2} = -\frac{3x^2}{y^4}(x^4 + y^4) = -\frac{3x^2}{y^4}$$

$$\frac{d^2y}{dx^2} = -\frac{3x^2}{y^7}$$

(ii)

$$\text{Width} = a_2 - a_1$$

$$\text{Area} = \int_{a_1}^{a_2} \frac{d^3y}{dx^3} dx = \left[\frac{d^2y}{dx^2}\right]_{a_1, b_1}^{a_2, b_2}$$

$$= \frac{3a_1^2}{b_1^7} - \frac{3a_2^2}{b_2^7}$$

$$= \frac{3a_1^2 b_2^7 - 3a_2^2 b_1^7}{b_1^7 b_2^7}$$

$$y_m = \frac{\text{Area}}{\text{Width}} = \frac{3(a_1^2 b_2^7 - a_2^2 b_1^7)}{b_1^7 b_2^7 (a_2 - a_1)}$$

Example 2: W2007 P1 Q8

Let $I_n = \int_0^{\ln 2} (e^x + e^{-x})^n \, dx$.

(i) Show that

$$\frac{d}{dx}\left[(e^x - e^{-x})(e^x + e^{-x})^{n-1}\right] = n(e^x + e^{-x})^n - 4(n-1)(e^x + e^{-x})^{n-2}.$$ [3]

(ii) Hence show that

$$nI_n = 4(n-1)I_{n-2} + \tfrac{3}{2}\left(\tfrac{5}{2}\right)^{n-1}.$$ [2]

(iii) Use the result in part **(ii)** to find the y-coordinate of the centroid of the region bounded by the axes, the line $x = \ln 2$ and the curve

$$y = (e^x + e^{-x})^2.$$

Give your answer correct to 3 decimal places. [5]

(i) $\dfrac{d}{dx}\left[\left(e^x-e^{-x}\right)\left(e^x+e^{-x}\right)^{n-1}\right]=\left(e^x+e^{-x}\right)\left(e^x+e^{-x}\right)^{n-1}+(n-1)\left(e^x-e^{-x}\right)^2\left(e^x+e^{-x}\right)^{n-2}$

$=\left(e^x+e^{-x}\right)^n+(n-1)\left[\left(e^x+e^{-x}\right)^2-4\right]\left(e^x+e^{-x}\right)^{n-2}$

$=\left(e^x+e^{-x}\right)^n+(n-1)\left(e^x+e^{-x}\right)^n-4(n-1)\left(e^x+e^{-x}\right)^{n-2}$

$=n\left(e^x+e^{-x}\right)^n-4(n-1)\left(e^x+e^{-x}\right)^{n-2}$

(ii) Integrating both sides in the interval $[0,\ln 2]$:

$\left[\left(e^x-e^{-x}\right)\left(e^x+e^{-x}\right)^{n-1}\right]_0^{\ln 2}=n\int_0^{\ln 2}\left(e^x+e^{-x}\right)^n dx-4(n-1)\int_0^{\ln 2}\left(e^x+e^{-x}\right)^{n-2}dx$

$\left(2-\dfrac{1}{2}\right)\left(2+\dfrac{1}{2}\right)^{n-1}=nI_n-4(n-1)I_{n-2}$

$nI_n=4(n-1)I_{n-2}+\left(\dfrac{3}{2}\right)\left(\dfrac{5}{2}\right)^{n-1}$

(iii) $n=0:\quad I_0=\displaystyle\int_0^{\ln 2}dx=\ln 2$ $\qquad$ $\text{Area}=\displaystyle\int_0^{\ln 2}(e^x+e^{-x})^2\,dx=I_2$

$n=2: I_2=2I_0+\dfrac{15}{8}=2\ln 2+\dfrac{15}{8}$ $\qquad$ $M_y=\dfrac{1}{2}\displaystyle\int_0^{\ln 2}(e^x+e^{-x})^4\,dx=\dfrac{1}{2}I_4$

$n=4: I_4=3I_2+\dfrac{375}{64}=6\ln 2+\dfrac{735}{64}$ $\qquad$ $\bar{y}=\dfrac{M_y}{Area}=\dfrac{I_4}{2I_2}=\dfrac{6\ln 2+\frac{735}{64}}{4\ln 2+\frac{15}{4}}\approx 2.398$

Example 3: S2011 P11 Q10

Let

$$I_n=\int_0^{\frac{1}{2}\pi}\cos^n x\,\mathrm{d}x,$$

where $n\geqslant 0$. Show that, for all $n\geqslant 2$,

$$I_n=\frac{n-1}{n}I_{n-2}.$$ [4]

A curve has parametric equations $x=a\sin^3 t$ and $y=a\cos^3 t$, where a is a constant and $0\leqslant t\leqslant \frac{1}{2}\pi$. Show that the mean value m of y over the interval $0\leqslant x\leqslant a$ is given by

$$m=3a\int_0^{\frac{1}{2}\pi}(\cos^4 t-\cos^6 t)\,\mathrm{d}t.$$ [4]

Find the exact value of m, in terms of a. [4]

$I_n=\int_0^{\frac{1}{2}\pi}\cos x.\cos^{n-1}x\,dx$

$=\left[\sin x\cos^{n-1}x\right]_0^{\frac{1}{2}\pi}-\int_0^{\frac{1}{2}\pi}\sin x.(n-1)\cos^{n-2}x(-\sin x)dx$

$=(n-1)\int_0^{\frac{1}{2}\pi}\sin^2 x\cos^{n-2}x\,dx$

$=(n-1)\left[\int_0^{\frac{1}{2}\pi}\cos^{n-2}x-\cos^n x\,dx\right]=(n-1)\left(I_{n-2}-I_n\right)$

$(1+n-1)I_n=(n-1)I_{n-2}\Rightarrow I_n=\dfrac{n-1}{n}I_{n-2}$

$$\text{Width} = a - 0 = a$$
$$x = a\sin^3 t$$
$$dx = 3a\sin^2 t\cos t\,dt$$

$$m = \frac{\text{Area}}{\text{Width}} = \frac{1}{a}\int_0^a y\,dx$$
$$= 3\int_0^{\frac{1}{2}\pi}\left(a\cos^3 t\right)\left(\sin^2 t\cos t\right)dt$$
$$= 3a\int_0^{\frac{1}{2}\pi}\left(\cos^4 t\right)\left(\sin^2 t\right)dt = 3a\int_0^{\frac{1}{2}\pi}\cos^4 t - \cos^6 t\,dt$$

$$n = 0: \quad I_0 = \int_0^{\frac{1}{2}\pi} dx = \frac{1}{2}\pi$$
$$n = 2: \quad I_2 = \frac{1}{2}I_0 = \frac{1}{4}\pi$$
$$n = 4: \quad I_4 = \frac{3}{4}I_2 = \frac{3}{16}\pi$$
$$n = 6: \quad I_6 = \frac{5}{6}I_4 = \frac{5}{32}\pi$$

$$m = 3a\int_0^{\frac{1}{2}\pi}\cos^4 t - \cos^6 t\,dt$$
$$= 3a\int_0^{\frac{1}{2}\pi}\cos^4 x - \cos^6 x\,dx$$
$$= 3a\left(I_4 - I_6\right)$$
$$= 3a\left(\frac{3}{16}\pi - \frac{5}{32}\pi\right) = \frac{3}{32}a\pi$$

Example 4: S2003 P1 Q11 EITHER

Given that

$$I_n = \int_0^1 x^n \mathrm{e}^{-\alpha x}\,\mathrm{d}x,$$

where α is a positive constant and n is a non-negative integer, show that for $n \geqslant 1$,

$$\alpha I_n = nI_{n-1} - \mathrm{e}^{-\alpha}. \qquad [3]$$

Hence, or otherwise, find the coordinates of the centroid of the finite region bounded by the x-axis, the line $x = 1$ and the curve $y = x\mathrm{e}^{-x}$, giving your answers in terms of e. [11]

$$I_n = \int_0^1 x^n e^{-\alpha x}dx$$
$$= \left[x^n.-\frac{1}{\alpha}e^{-\alpha x}\right]_0^1 + \frac{1}{\alpha}\int_0^1 e^{-\alpha x}.nx^{n-1}dx$$
$$= -\frac{1}{\alpha}e^{-\alpha} + \frac{n}{\alpha}\int_0^1 x^{n-1}e^{-\alpha x}dx$$
$$\alpha I_n = nI_{n-1} - e^{-\alpha}$$

$$n = 0: \quad I_0 = \int_0^1 e^{-\alpha x}\,dx$$
$$= -\frac{1}{\alpha}[e^{-\alpha x}]_0^1 = \frac{1}{\alpha}(1 - e^{-\alpha})$$
$$n = 1: I_1 = \frac{1}{\alpha}I_0 - \frac{1}{\alpha}e^{-\alpha}$$
$$= \frac{1}{\alpha^2}(1 - e^{-\alpha}) - \frac{1}{\alpha}e^{-\alpha}$$
$$n = 2: I_2 = \frac{2}{\alpha}I_1 - \frac{1}{\alpha}e^{-\alpha}$$
$$= \frac{2}{\alpha^3}(1 - e^{-\alpha}) - \frac{2}{\alpha^2}e^{-\alpha} - \frac{1}{\alpha}e^{-\alpha}$$

$$\text{Area} = \int_0^1 xe^{-x}\,dx = I_1, \alpha = 1$$
$$= 1 - 2e^{-1}$$
$$M_x = \int_0^1 x(xe^{-x})\,dx = I_2, \alpha = 1$$
$$= 2 - 5e^{-1}$$
$$M_y = \int_0^1 \frac{1}{2}(xe^{-x})^2\,dx = \frac{1}{2}I_2, \alpha = 2$$
$$= \frac{1}{8} - \frac{5}{8}e^{-2}$$

$$(\bar{x}, \bar{y}) = \left(\frac{M_x}{Area}, \frac{M_y}{Area}\right)$$
$$= \left(\frac{2 - 5e^{-1}}{1 - 2e^{-1}}, \frac{1 - 5e^{-2}}{8\left(1 - 2e^{-1}\right)}\right)$$
$$= \left[\frac{2e - 5}{e - 2}, \frac{e^2 - 5}{8e(e - 2)}\right]$$

Example 5: W2006 P1 Q12 OR

The curve C has equation

$$y = x^{\frac{1}{2}} - \tfrac{1}{3}x^{\frac{3}{2}} + \lambda,$$

where $\lambda > 0$ and $0 \leqslant x \leqslant 3$. The length of C is denoted by s. Prove that $s = 2\sqrt{3}$. [4]

The area of the surface generated when C is rotated through one revolution about the x-axis is denoted by S. Find S in terms of λ. [5]

The y-coordinate of the centroid of the region bounded by C, the axes and the line $x = 3$ is denoted by h. Given that $\int_0^3 y^2\,dx = \frac{3}{4} + \frac{8\sqrt{3}}{5}\lambda + 3\lambda^2$, show that

$$\lim_{\lambda\to\infty} \frac{S}{hs} = 4\pi. \quad [5]$$

$$y = x^{\frac{1}{2}} - \frac{1}{3}x^{\frac{3}{2}} + \lambda \rightarrow \frac{dy}{dx} = \frac{1}{2}x^{-\frac{1}{2}} - \frac{1}{2}x^{\frac{1}{2}}$$

$$\sqrt{1+\left(\frac{dy}{dx}\right)^2} = 1 + \left(\frac{1}{2}x^{-\frac{1}{2}} - \frac{1}{2}x^{\frac{1}{2}}\right)^2 = \left(\frac{1}{2}x^{-\frac{1}{2}} + \frac{1}{2}x^{\frac{1}{2}}\right)^2$$

$$s = \int_0^3 \sqrt{1+\left(\frac{dy}{dx}\right)^2}\,dx = \frac{1}{2}\int_0^3 x^{-0.5} + x^{0.5}\,dx = \frac{1}{2}\left[2x^{0.5} + \frac{2}{3}x^{1.5}\right]_0^3 = \frac{1}{2}\left(2\sqrt{3} + 2\sqrt{3}\right) = 2\sqrt{3}$$

$$S = \int_0^3 2\pi y\sqrt{1+\left(\frac{dy}{dx}\right)^2}\,dx = \pi\int_0^3 \left(x^{0.5} - \frac{1}{3}x^{1.5} + \lambda\right)\left(x^{-0.5} + x^{0.5}\right)dx = \pi\int_0^3 1 - \frac{1}{3}x + \lambda x^{-0.5} + x - \frac{1}{3}x^2 + \lambda x^{0.5}\,dx$$

$$= \pi\left[x + \frac{1}{3}x^2 + \frac{2}{3}\lambda x^{1.5} + 2\lambda x^{0.5} - \frac{1}{9}x^3\right]_0^3 = \pi\left(3 + 3 + 2\lambda\sqrt{3} + 2\lambda x\sqrt{3} - 3\right) = \pi\left(3 + 4\lambda\sqrt{3}\right)$$

$$Area = \int_0^3 x^{0.5} - \frac{1}{3}x^{1.5} + \lambda\,dx = \left[\frac{2}{3}x^{1.5} - \frac{2}{15}x^{2.5} + \lambda x\right]_0^3 = 2\sqrt{3} - \frac{18}{15}\sqrt{3} + 3\lambda = 3\lambda + \frac{4}{5}\sqrt{3}$$

$$h = \frac{M_y}{Area} = \frac{\frac{3}{4} + \frac{8}{5}\sqrt{3}\lambda + 3\lambda^2}{6\lambda + \frac{8}{5}\sqrt{3}}$$

$$\frac{S}{hs} = \frac{\pi\left(3 + 4\lambda\sqrt{3}\right)\left(6\lambda + \frac{8}{5}\sqrt{3}\right)}{2\sqrt{3}\left(\frac{3}{4} + \frac{8}{5}\sqrt{3}\lambda + 3\lambda^2\right)} = \frac{\pi\left(18\lambda + 24\lambda^2\sqrt{3} + \frac{96}{5}\lambda + \frac{24}{5}\sqrt{3}\right)}{\frac{3}{2}\sqrt{3} + \frac{48}{5}\lambda + 6\sqrt{3}\lambda^2}$$

$$= \frac{\pi\left(18\lambda^{-1} + 24\sqrt{3} + \frac{96}{5}\lambda^{-1} + \frac{24}{5}\sqrt{3}\lambda^{-2}\right)}{\frac{3}{2}\sqrt{3}\lambda^{-2} + \frac{48}{5}\lambda^{-1} + 6\sqrt{3}}$$

$$\lim_{\lambda\to\infty}\frac{S}{hs} = \frac{\pi\left(24\sqrt{3}\right)}{6\sqrt{3}} = 4\pi$$

CHAPTER 6: DIFFERENTIAL EQUATIONS

6.1 Second Order Linear Differential Equations

- LO: recall the meaning of the terms 'complementary function' and 'particular integral' in the context of linear differential equations, and use the fact that the general solution is the sum of the complementary function and a particular integral;
- LO: find the complementary function for a first or second order linear differential equation with constant coefficients;
- LO: recall the form of, and find, a particular integral for a first or second order linear differential equation in the cases where $ax^2 + bx + c$ or ae^{bx} or $a\cos px + b\sin px$ is a suitable form, and in other cases find the appropriate coefficient(s) given a suitable form of particular integral;
- LO: use initial conditions to find a particular solution to a differential equation, and interpret the solution in the context of a problem modelled by the differential equation;

A) Pre-Requisite Knowledge (Maths Paper 3 Differential Equations)

- Definitions: differential equation (DE), order of DE, power of DE,
- general solutions, particular solutions, and boundary conditions

B) Second Order Homogenous Linear Differential Equations

- Second Order Linear Differential Equations are in the form $a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f(x)$. The equation is second order because the term $\frac{d^2y}{dx^2}$ is present.
- When $f(x) = 0$ the equation is called homogenous.
- In this section, we will investigate the homogenous case of this equation:

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0$$

- The following is a complete proof for the solution to this differential equation (unobtainable in most course books). Still I will post a summary on how to solve these in the end. The method in the summary is sufficient for full points in the exam.
- To master this proof, you are recommended to have studied first order linear differential equations (e.g. how to solve the equation $\frac{dy}{dx} + P(x)y = Q(x)$ using an integrating factor) although that is not in the syllabus.

Solution: Let the equation $am^2 + bm + c = 0$ have roots $m = \alpha$ and $m = \beta$. Thus

$$\alpha + \beta = -\frac{b}{a}, \quad \alpha\beta = \frac{c}{a}$$

We can now rewrite the original equation into:

$$\frac{d^2y}{dx^2} + \frac{b}{a}\frac{dy}{dx} + \frac{c}{a}y = 0$$

$$\frac{d^2y}{dx^2} - \alpha\frac{dy}{dx} - \beta\frac{dy}{dx} + \alpha\beta y = 0$$

Now we break the left hand side into two parts and investigate them.

(1): Consider $-\beta\frac{dy}{dx}+\alpha\beta y$.

Multiply the expression by $f(x)$ and let

$$-\beta\left(f(x)\frac{dy}{dx}-\alpha f(x)y\right)=-\beta\frac{d}{dx}[yf(x)]$$
$$=-\beta\left[f(x)\frac{dy}{dx}+f'(x)y\right]$$

Comparing coefficients,

$$f'(x)=-\alpha f(x)\rightarrow\frac{f'(x)}{f(x)}=-\alpha$$

Integrating both sides:

$$\ln f(x)=-\alpha x\rightarrow f(x)=e^{-\alpha x}$$

(2): Consider $\frac{d^2y}{dx^2}-\alpha\frac{dy}{dx}$.

Multiply the expression by $g(x)$ and let

$$g(x)\frac{d^2y}{dx^2}-\alpha g(x)\frac{dy}{dx}=\frac{d}{dx}\left[g(x)\frac{dy}{dx}\right]$$
$$=g(x)\frac{d^2y}{dx^2}+g'(x)\frac{dy}{dx}$$

Comparing coefficients,

$$g'(x)=-\alpha g(x)\rightarrow\frac{g'(x)}{g(x)}=-\alpha$$

Similarly to the left, $g(x)=e^{-\alpha x}$.

Rewriting and integrating both sides (we only multiply RHS with $f(x)$ as $f(x)=g(x)$):

$$\frac{d}{dx}\left[e^{-\alpha x}\frac{dy}{dx}\right]-\beta\frac{d}{dx}[e^{-\alpha x}y]=0.f(x)=0$$
$$e^{-\alpha x}\frac{dy}{dx}-\beta e^{-\alpha x}y=c\rightarrow\ \frac{dy}{dx}+\beta y=ce^{\alpha x}$$

Once again multiply the equation by $h(x)$ and let

$$h(x)\frac{dy}{dx}-\beta h(x)y=\frac{d}{dx}[yh(x)]=h(x)\frac{dy}{dx}+h'(x)y$$

Comparing the coefficients we have $h'(x)=-\beta h(x)$ so $h(x)=e^{-\beta x}$ as shown above.

$$\frac{d}{dx}\left[ye^{-\beta x}\right]=ce^{\alpha x}e^{-\beta x}=ce^{(\alpha-\beta)x}-(2)$$

Case 1: When $\alpha\neq\beta$ (the equation has two distinct real roots)

integrating both sides give

$$ye^{-\beta x}=\frac{c}{\alpha-\beta}e^{(\alpha-\beta)x}+d\rightarrow y=\frac{c}{\alpha-\beta}e^{\alpha x}+de^{\beta x}-(3)$$

The general solution is $y=Ae^{\alpha x}+Be^{\beta x}$, where $A=\frac{c}{\alpha-\beta},B=d$.

Case 2: When $\alpha=\beta$ (the equation has two equal roots)

(2) gets reduced to

$$\frac{d}{dx}\left[ye^{-\beta x}\right]=ce^{0x}=c$$

Integrating gives $ye^{-\beta x}=cx+d\rightarrow y=(cx+d)e^{\beta x}$

The general solution is $y=(A+Bx)e^{\alpha x}$, where $A=d,B=c$ and $\alpha=\beta$.

Case 3: When the equation has two complex roots,

let $\alpha=p+qi,\beta=p-qi$:

$$e^{\alpha x}=e^{px+qix}=e^{px}e^{qix}=e^{px}(\cos qx+i\sin qx)$$
$$e^{\beta x}=e^{px-qix}=e^{px}e^{-qix}=e^{px}(\cos qx-i\sin qx)$$

Substituting into (3) gives:

$$y=\frac{c}{2qi}e^{px}(\cos qx+i\sin qx)+de^{px}(\cos qx-i\sin qx)$$

The general solution is $y=e^{px}(A\cos qx+B\sin qx)$, where $A=\frac{c}{2qi}+d,B=i\left(\frac{c}{2qi}-d\right)$.

Note: When finding $f(x)$, $g(x)$ and $h(x)$, a constant is not added during integration as it will just be cancelled off when we multiply the entire equation with the function.

Summary: Obtain the auxiliary equation $am^2 + bm + c = 0$ from the equation

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0.$$

1. If the auxiliary equation has two distinct real roots α and β, then $y = Ae^{\alpha x} + Be^{\beta x}$.

2. If the auxiliary equation has two equal real roots α, then $y = (A + Bx)e^{\alpha x}$.

3. If the auxiliary equation has two complex roots $p \pm qi$, then $y = e^{px}(A\cos qx + B\sin qx)$.

This solution is known as the complementary function.

Example: Find the general solution of the following differential equations:

1. $\frac{d^2y}{dx^2} + 5\frac{dy}{dx} + 6y = 0$ Consider $m^2 + 5m + 6 = 0$ $(m + 2)(m + 3) = 0$ $m = -2, -3.$ GS: $y = Ae^{-2x} + Be^{-3x}$	2. $\frac{d^2y}{dx^2} + \frac{dy}{dx} - 12y = 0$ Consider $m^2 + m - 12 = 0$ $(m + 4)(m - 3) = 0$ $m = 3, -4.$ GS: $y = Ae^{3x} + Be^{-4x}$
3. $\frac{d^2y}{dx^2} - 16\frac{dy}{dx} + 64y = 0$ Consider $m^2 - 16m + 64 = 0$ $(m - 8)(m - 8) = 0$ $m = 8$, GS: $y = (A + Bx)e^{8x}$	4. $4\frac{d^2y}{dx^2} + 20\frac{dy}{dx} + 25y = 0$ Consider $4m^2 + 20m + 25 = 0$ $(2m + 5)(2m + 5) = 0$ $m = -\frac{5}{2}$, GS: $y = (A + Bx)e^{-\frac{5}{2}x}$
5. $\frac{d^2y}{dx^2} + y = 0$ Consider $m^2 + 1 = 0$ $m = i, -i.$ GS: $y = A\cos x + B\sin x$ Note: This is the equation for simple harmonic motion in mechanics (Topic 13).	6. $\frac{d^2y}{dx^2} + 8\frac{dy}{dx} + 17y = 0$ Consider $m^2 + 8m + 17 = 0$ $m = \frac{-8 \pm \sqrt{64 - 68}}{2} = -4 \pm i$ GS: $y = e^{-4x}(A\cos x + B\sin x)$

C) Second Order Non-Homogenous Linear Differential Equations

In this section, we extend the idea from the previous section to solve the differential equation

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f(x)$$

1. Method of Undetermined Coefficients

- Using this method, a suitable function with unknown constants is substituted into the differential equation to determine the values of the coefficients.
- This technique only works $f(x)$ can only be one of the following types of functions;

a) a polynomial of degree 2 or less, $ax^2 + bx + c$;
b) an exponential function, ae^{bx};
c) a trigonometric function in terms of $\sin x$ and / or $\cos x$, $a\cos px + b\sin px$.

- The table below shows the function to be substituted:

When $f(x)$ is a …	Substitute $y_p =$
Polynomial Example: 1. $p + qx + rx^2$ 2. $p + qx$	A polynomial of the same degree Try: $C + Dx + Ex^2$ $C + Dx$
An exponential function Example: 1. e^{6x} 2. $2e^{-2x}$	The same function times a constant Try: Ce^{6x} Ce^{-2x}
A trigonometric function Example: 1. $A\cos 2x + B\sin 2x$ 2. $A\sin 5x$ 3. $B\cos x$	A combination of sine and cosine Try: $C\cos 2x + D\sin 2x$ $C\cos 5x + D\sin 5x$ $C\cos x + D\sin x$ Note: Always use a combination and cosine and sine even if only one of them is in $f(x)$.

- Main technique: choose the appropriate y, differentiate for $\frac{dy}{dx}$ and $\frac{d^2y}{dx^2}$, substitute into the differentiation and compare the coefficients to find the constants.
- For other functions, a different method known as the method of variation of parameters is used (not in syllabus)
- If the function tried is present in the complementary function, try that function times by x or x^2 (this is rare, but if it comes out in the exam a hint will usually be given)

Example: Find a particular integral for the following differential equations:

1.

$$\frac{d^2y}{dx^2} - 4\frac{dy}{dx} + 5y = 25x^2 - 7$$

Try $y = cx^2 + dx + e \rightarrow \frac{dy}{dx} = 2cx + d \rightarrow \frac{d^2y}{dx^2} = 2c$

Substitute into the equation,

$$2c + 4(2cx + d) + 5(cx^2 + dx + e) = 25x^2 - 7$$

$$(2c + 4d + 5e) + x(8c + 5d) + 5cx^2 = 25x^2 - 7$$

Comparing coefficients: x^2: $5c = 25,\ c = 5$

$$x\text{: } 8(5) + 5d = 0,\ d = -8$$

$$2(5) + 4(-8) + 5e = -7,\ e = 3$$

Particular integral is $y = 5x^2 - 8x + 3$.

2.

$$\frac{d^2y}{dx^2} + \frac{dy}{dx} - 12y = 12e^{2x}$$

Try $y = ce^{2x} \rightarrow \frac{dy}{dx} = 2ce^{2x} \rightarrow \frac{d^2y}{dx^2} = 4ce^{2x}$

$$4ce^{2x} + 2ce^{2x} - 12(ce^{2x}) = 12e^{2x}$$

$$-6ce^{2x} = 12e^{2x} \rightarrow c = -2$$

Particular integral is $y = -2e^{2x}$.

3.	$$\frac{d^2y}{dx^2} + 2\frac{dy}{dx} + y = 25\cos 2x$$ $y = C\cos 2x + D\sin 2x \rightarrow \frac{dy}{dx} = 2D\cos 2x - 2C\sin 2x \rightarrow \frac{d^2y}{dx^2} = -4C\cos 2x - 4D\sin 2x$ $(-4C\cos 2x - 4D\sin 2x) + 2(2D\cos 2x - 2C\sin 2x) + C\cos 2x + D\sin 2x = 25\cos 2x$ $(4D - 3C)\cos 2x + (-3D - 4C)\sin 2x = 25\cos 2x$ Comparing coefficients, $4D - 3C = 25 - (1), \quad 3D + 4C = 0 - (2)$ Solving gives $D = 4, C = -3$, particular integral is $y = 4\sin 2x - 3\cos 2x$.
4.	$$\frac{d^2y}{dx^2} - y = 2e^x$$ As e^x is present in the complementary function, try $y = cxe^x$. $$y = cxe^x \rightarrow \frac{dy}{dx} = (cx + c)e^x \rightarrow \frac{d^2y}{dx^2} = (cx + 2c)e^x$$ $$(cx + 2c)e^x - cxe^x = 2e^x \rightarrow 2ce^x = 2e^x.$$ Comparing coefficients, $c = 1$, thus particular integral is $y = xe^x$.
5.	$$\frac{d^2y}{dx^2} - 4\frac{dy}{dx} + 4y = 4e^{2x}$$ As e^{2x} and xe^{2x} are present in the complementary function, try $y = cx^2e^{2x}$. $$y = cx^2e^{2x} \rightarrow \frac{dy}{dx} = (2cx + 2cx^2)e^{2x} \rightarrow \frac{d^2y}{dx^2} = (2c + 8cx + 4cx^2)e^{2x}$$ $$(2c + 8cx + 4cx^2)e^{2x} - 4(2cx + 2cx^2)e^{2x} + 4cx^2e^{2x} = 4e^{2x}$$ $$2ce^{2x} = 4e^{2x} \rightarrow c = 2$$ Particular integral is $y = 2x^2e^{2x}$.

2. Principle of superposition:

> If $P(x)$ is a solution to $a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0$,
> and $Q(x)$ is a solution to $a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f(x)$,
> then $P(x) + Q(x)$ is also a solution to $a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f(x)$.

Proof:

$P(x)$ is a solution to $a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0 \rightarrow aP''(x) + bP'(x) + cP(x) = 0 - (1)$

$Q(x)$ is a solution to $a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f(x) \rightarrow aQ''(x) + bQ'(x) + cQ(x) = f(x) - (2)$

$(1) + (2)$: $\quad a[P''(x) + Q''(x)] + b[P'(x) + Q'(x)] + c[P(x) + Q(x)] = f(x)$

Thus a solution is $y = P(x) + Q(x)$.

Referring to the above:

- $P(x)$ is known as the complementary function, y_c.
- $Q(x)$ is known as the particular integral, y_p.
- Thus the general solution is $y = y_c + y_p$.

D) Particular Solutions

- When boundary conditions are given a particular solution can be obtained from a general solution.
- For a second-order equation, TWO conditions are needed to obtain a particular solution, as opposed to only one for first-order equations in Maths P3.
- Typically, either one of the two conditions will be given:

1. The value of y for two different values of x;

2. The value of y and the value of $\frac{dy}{dx}$ for a particular value of x. (More popular in exams)

Steps to solve exam questions:
1. Find the complementary function
(take auxiliary QE from the DE and use one of the 3 forms based on the type of roots)
2. Find the particular integral
(choose the appropriate form to use: trigonometry / exponential / polynomial)
3. State the general solution GS = complementary function + particular integral
4. Differentiate and apply initial conditions to find the particular solution.

Example 1: W2002 P1 Q8

The value of the assets of a large commercial organisation at time t, measured in years, is $\$(10^8 y + 10^9)$. The variables y and t are related by the differential equation

$$\frac{d^2y}{dt^2} + 5\frac{dy}{dt} + 6y = 15\cos 3t - 3\sin 3t.$$

Find y in terms of t, given that $y = 3$ and $\frac{dy}{dt} = -2$ when $t = 0$. [9]

Show that, for large values of t, the value of the assets is less than $\$9.5 \times 10^8$ for about a third of the time. [3]

Auxiliary equation is $m^2 + 5m + 6 = 0 \rightarrow (m+2)(m+3) = 0 \rightarrow m = -2, -3$

Hence the complementary function, $y_c = Ae^{-3t} + Be^{-2t}$.

Let the particular integral, $y_p = C\cos 3t + D\sin 3t$.

Hence $\frac{d}{dt}(y_p) = 3D\cos 3t - 3C\sin 3t$, $\frac{d^2}{dt^2}(y_p) = -9C\cos 3t - 9D\sin 3t$.

Substitute into the original differential equation:

$$\frac{d^2}{dt^2}(y_p) + 5\frac{d}{dt}(y_p) + 6y_p = 15\cos 3t - 3\sin 3t$$

$$15D\cos 2t - 15C\sin 2t - 3C\cos 2t - 3D\sin 2t = 15\cos 3t - 3\sin 3t$$

Comparing the coefficients,

$$15D - 3C = 15, -15C - 3D = -3 \rightarrow C = 0, D = 1.$$

Hence $y_p = 3\sin 2t$. The general solution is then

$$y = y_c + y_p = Ae^{-3t} + Be^{-2t} + \sin 3t$$

$$t = 0, y = 3 \rightarrow A + B = 3 - (1)$$

$$\frac{dy}{dt} = -3Ae^{-3t} - 2Be^{-2t} + 3\cos 3t$$

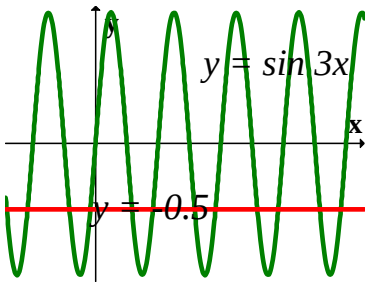
$$t = 0, \frac{dy}{dt} = -2 \rightarrow -3A - 2B + 3 = -2 - (2)$$

Solving (1) and (2) gives $A = -1, B = 4$. Hence

$$y = 4e^{-2t} - e^{-3t} + \sin 3t$$

When $\rightarrow \infty$, $e^{-2t} \rightarrow 0, e^{-3t} \rightarrow 0$, hence $y \rightarrow \sin 3t$.

Assets less than $9.5 \times 10^8 \rightarrow y < -0.5$.



Analyse a period of $\frac{2\pi}{3}$ (one revolution)

When $y = -\frac{1}{2}$, $x = \frac{7}{18}\pi, \frac{11}{18}\pi, \ldots$

Interval width = $\frac{11}{18}\pi - \frac{7}{18}\pi = \frac{2}{9}\pi$.

Since interval width is $\frac{1}{3}$ of the period, the assets are less than $9.5 \times 10^8 \frac{1}{3}$ of the time.

Example 2: W2003 P1 Q7

Find the general solution of the differential equation

$$\frac{d^2y}{dt^2} + 4\frac{dy}{dt} + 4y = e^{-\alpha t},$$

where α is a constant and $\alpha \neq 2$. [7]

Show that if $\alpha < 2$ then, whatever the initial conditions, $ye^{\alpha t} \rightarrow \frac{1}{(2-\alpha)^2}$ as $t \rightarrow \infty$. [2]

Auxiliary equation is $m^2 + 4m + 4 = 0 \rightarrow (m + 2) = 0 \rightarrow m = -2$

Hence the complementary function, $y_c = (A + Bt)e^{-2t}$.

Let the particular integral, $y_p = Ce^{-\alpha t}$. Hence $\frac{d}{dx}(y_p) = -\alpha Ce^{-\alpha t}$, $\frac{d^2}{dx^2}(y_p) = \alpha^2 Ce^{-\alpha t}$.

Substitute into the original differential equation:

$$\frac{d^2}{dx^2}(y_p) + 4\frac{d}{dx}(y_p) + 4y_p = e^{-\alpha t}$$

$$\alpha^2 Ce^{-\alpha t} - 4\alpha Ce^{-\alpha t} + Ce^{-\alpha t} = e^{-\alpha t}$$

$$C(\alpha^2 - 4\alpha + 4) = 1 \rightarrow C = \frac{1}{(\alpha - 2)^2}$$

Hence $y_p = \frac{1}{(\alpha-2)^2}e^{-\alpha t}$. The general solution is then

$$y = y_c + y_p = (A + Bt)e^{-2t} + \frac{1}{(\alpha - 2)^2}e^{-\alpha t}$$

$$ye^{\alpha t} = (A + Bt)e^{(\alpha-2)t} + \frac{1}{(\alpha - 2)^2}$$

If $\alpha < 2, \alpha - 2 < 0$. When $t \to \infty$, $e^{(\alpha-2)t} \to 0$, hence $ye^{\alpha t} \to \frac{1}{(\alpha-2)^2}$.

Example 3: W2008 P1 Q8

Find y in terms of t, given that

$$5\frac{\mathrm{d}^2y}{\mathrm{d}t^2} + 6\frac{\mathrm{d}y}{\mathrm{d}t} + 5y = 15 + 12t + 5t^2,$$

and that $y = \frac{\mathrm{d}y}{\mathrm{d}t} = 0$ when $t = 0$. [9]

Auxiliary equation is $5m^2 + 6m + 5 = 0 \to m = \frac{-6\pm\sqrt{6^2-4(25)}}{2(5)} = -\frac{3}{5} \pm \frac{4}{5}i$.

Hence the complementary function, $y_c = e^{-\frac{3}{5}t}\left(A\cos\frac{4}{5}t + B\sin\frac{4}{5}t\right)$

Let the particular integral, $y_p = C + Dt + Et^2$. Hence $\frac{d}{dt}(y_p) = D + 2Et$, $\frac{d^2}{dt^2}(y_p) = 2E$.

Substitute into the original differential equation:

$$5\frac{d^2}{dt^2}(y_p) + 6\frac{d}{dt}(y_p) + 5y_p = 15 + 12t + 5t^2$$

$$5Et^2 + t(12E + 5D) + (10E + 6D + 5C) = 15 + 12t + 5t^2$$

Comparing the coefficients,

$$5E = 5 \to E = 1, 12E + 5D = 12 \to D = 0, 10E + 5C + 6D = 15, C = 1.$$

Hence $y_p = 1 + t^2$. The general solution is then

$$y = y_c + y_p = e^{-\frac{3}{5}t}\left(A\cos\frac{4}{5}t + B\sin\frac{4}{5}t\right) + 1 + t^2$$

$$t = 0, y = 0 \to \quad A(1) + 1 = 0 \to A = -1$$

$$\frac{dy}{dt} = -\frac{3}{5}e^{-\frac{3}{5}t}\left(A\cos\frac{4}{5}t + B\sin\frac{4}{5}t\right) + e^{-\frac{3}{5}t}\left(\frac{4}{5}B\cos\frac{4}{5}t - \frac{4}{5}A\sin\frac{4}{5}t\right) + 2t$$

$$t = 0, \frac{dy}{dt} = 0 \to \quad -\frac{3}{5}A + \frac{4}{5}B = 0 \to B = -\frac{3}{4}$$

$$\therefore y = -\frac{1}{4}e^{-\frac{3}{5}t}\left(4\cos\frac{4}{5}t + 3\sin\frac{4}{5}t\right) + 1 + t^2$$

6.2 Reducing Other Differential Equations to Second Order

- LO: use a given substitution to reduce a given differential equation to a second order linear equation with constant coefficients;

  - Note: You are advised to finish Section 5.1 before starting this part.
  - Substitutions can be used to reduce other differential equations to ones that are solvable with methods seen in the previous section.
  - General strategy: Differentiate the given substitution with respect to a suitable variable twice, then replace the related terms to obtain an easier equation to solve.

Example 1: W2009 P1 Q9

Show that if y depends on x and $x = e^u$ then

$$x^2\frac{d^2y}{dx^2} = \frac{d^2y}{du^2} - \frac{dy}{du}. \qquad [4]$$

Given that y satisfies the differential equation

$$x^2\frac{d^2y}{dx^2} + 5x\frac{dy}{dx} + 3y = 30x^2,$$

use the substitution $x = e^u$ to show that

$$\frac{d^2y}{du^2} + 4\frac{dy}{du} + 3y = 30e^{2u}. \qquad [2]$$

Hence find the general solution for y in terms of x. [5]

$$x = e^u \rightarrow \frac{dx}{du} = e^u = x$$

$$\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx} = \frac{1}{x}\frac{dy}{du}$$

$$\frac{d^2y}{dx^2} = \frac{1}{x}\frac{d}{dx}\left(\frac{dy}{du}\right) - \frac{1}{x^2}\frac{dy}{du}$$

$$x^2\frac{d^2y}{dx^2} = \frac{x\frac{d^2y}{du^2}}{\frac{dx}{du}} - \frac{dy}{du} \rightarrow x^2\frac{d^2y}{dx^2} = \frac{d^2y}{du^2} - \frac{dy}{du}$$

Auxiliary equation is $m^2 + 4m + 3 = 0 \rightarrow (m+3)(m+1) = 0 \rightarrow m = -1, -3$

Hence the complementary function, $y_c = Ae^{-3u} + Be^{-u}$.

Let the particular integral, $y_p = Ce^{2u}$. Hence $\frac{d}{du}(y_p) = 2Ce^{2u}$, $\frac{d^2}{du^2}(y_p) = 4Ce^{2u}$.

Substitute into the original differential equation:

$$\frac{d^2}{du^2}(y_p) + 4\frac{d}{du}(y_p) + 3y_p = 30e^{2u}$$

$$4Ce^{2u} + 8Ce^{2u} + 3Ce^{2u} = 30e^{2u}$$

$$15C = 30 \rightarrow C = 2$$

Hence $y_p = 2e^{2u}$. The general solution is then

$$y = y_c + y_p = Ae^{-3u} + Be^{-u} + 2e^{2u}$$

$$x = e^u \rightarrow y = \frac{A}{x^3} + \frac{B}{x} + 2x^2$$

Example 2: S2003 P1 Q9

The variables x and t, where $x > 0$ and $0 \leqslant t \leqslant \frac{1}{2}\pi$, are related by

$$x\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + \left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + 5x\frac{\mathrm{d}x}{\mathrm{d}t} + 3x^2 = 3\sin 2t + 15\cos 2t,$$

and the variables x and y are related by $y = x^2$. Show that

$$\frac{\mathrm{d}^2y}{\mathrm{d}t^2} + 5\frac{\mathrm{d}y}{\mathrm{d}t} + 6y = 6\sin 2t + 30\cos 2t. \qquad [3]$$

Hence find x in terms of t, given that $x = 2$ and $\frac{\mathrm{d}x}{\mathrm{d}t} = -\frac{3}{2}$ when $t = 0$. [10]

$$y = x^2 \rightarrow \frac{dy}{dt} = 2x\frac{dx}{dt} \rightarrow \frac{d^2y}{dt^2} = 2x\frac{d^2x}{dt^2} + 2\left(\frac{dx}{dt}\right)^2$$

$$x\frac{d^2x}{dt^2} + \left(\frac{dx}{dt}\right)^2 + 5x\frac{dx}{dt} + 3x^2 = 15\cos 2t + 3\sin 2t$$

$$\frac{1}{2}\frac{d^2y}{dt^2} + \frac{5}{2}\frac{dy}{dt} + 3y = 15\cos 2t + 3\sin 2t$$

$$\frac{d^2y}{dt^2} + 5\frac{dy}{dt} + 6y = 30\cos 2t + 6\sin 2t$$

Auxiliary equation is $m^2 + 5m + 6 = 0 \rightarrow (m+2)(m+3) = 0 \rightarrow m = -2, -3$

Hence the complementary function, $y_c = Ae^{-3t} + Be^{-2t}$.

Let the particular integral, $x_p = C\cos 2t + D\sin 2t$.

Hence $\frac{d}{dt}(y_p) = 2D\cos 2t - 2C\sin 2t$, $\frac{d^2}{dt^2}(y_p) = -4C\cos 2t - 4D\sin 2t$.

Substitute into the original differential equation:

$$\frac{d^2}{dt^2}(y_p) + 5\frac{d}{dt}(y_p) + 6y_p = 30\cos 2t + 6\sin 2t$$

$$10D\cos 2t - 10C\sin 2t + 2C\cos 2t + 2D\sin 2t = 30\cos 2t + 6\sin 2t$$

Comparing the coefficients,

$$10D + 2C = 30, -10C + 2D = 6 \rightarrow C = 0, D = 3.$$

Hence $x_p = 3\sin 2t$. The general solution is then

$$y = y_c + y_p = Ae^{-3t} + Be^{-2t} + 3\sin 2t$$

$$x^2 = Ae^{-3t} + Be^{-2t} + 3\sin 2t$$

$$t = 0, x = 2 \rightarrow A + B = 4 - (1)$$

$$2x\frac{dx}{dt} = -3Ae^{-3t} - 2Be^{-2t} + 6\cos 2t$$

$$t = 0, x = 2, \frac{dx}{dt} = -\frac{3}{2} \rightarrow -3A - 2B + 6 = -6 - (2)$$

Solving (1) and (2) gives $A = 4, B = 0$. Thus

$$x^2 = 4e^{-3t} + 3\sin 2t \rightarrow \quad x = \sqrt{4e^{-3t} + 3\sin 2t}$$

Example 3: W2007 P1 Q12 OR

Show that the substitution $y = \frac{1}{w}$ reduces the differential equation

$$y\frac{d^2y}{dx^2} + 2y\frac{dy}{dx} - 2\left(\frac{dy}{dx}\right)^2 - 5y^2 = (5x^2 + 4x + 2)y^3$$

to

$$\frac{d^2w}{dx^2} + 2\frac{dw}{dx} + 5w = -5x^2 - 4x - 2. \quad [4]$$

Find the general solution for w in terms of x. [6]

Find a function f such that $\lim_{x\to\infty}\left(\frac{y}{f(x)}\right) = 1$. [3]

$$y = \frac{1}{w} \to \frac{dy}{dx} = -\frac{1}{w^2}\frac{dw}{dx} \to \frac{d^2y}{dx^2} = \frac{2}{w^3}\left(\frac{dw}{dx}\right)^2 - \frac{1}{w^2}\frac{d^2w}{dx^2}$$

$$y\frac{d^2y}{dx^2} + 2y\frac{dy}{dx} - 2\left(\frac{dy}{dx}\right)^2 - 5y^2 = (5x^2 + 4x + 2)y^3$$

$$\frac{2}{w^4}\left(\frac{dw}{dx}\right)^2 - \frac{1}{w^3}\frac{d^2w}{dx^2} - \frac{2}{w^3}\frac{dw}{dx} - \frac{2}{w^4}\left(\frac{dw}{dx}\right)^2 - \frac{5}{w^2} = \frac{5x^2 + 4x + 2}{w^3}$$

$$-\frac{1}{w^3}\frac{d^2w}{dx^2} - \frac{2}{w^3}\frac{dw}{dx} - \frac{5}{w^2} = \frac{5x^2 + 4x + 2}{w^3}$$

$$\frac{d^2w}{dx^2} + 2\frac{dw}{dx} + 5w = -5x^2 - 4x - 2$$

Auxiliary equation is $m^2 + 2m + 5 = 0 \to m = \frac{-2\pm\sqrt{2^2-4(5)}}{2} = -1 \pm 2i$.

Hence the complementary function, $w_c = e^{-x}(A\cos 2x + B\sin 2x)$

Let particular integral, $w_p = C + Dx + Ex^2$. Hence $\frac{d}{dx}(w_p) = D + 2Ex$, $\frac{d^2}{dx^2}(w_p) = 2E$.

Substitute into the original differential equation:

$$\frac{d^2}{dx^2}(w_p) + 2\frac{d}{dx}(w_p) + 5w_p = -5x^2 - 4x - 2$$

$$5Ex^2 + x(4E + 5D) + (5C + 2D + 2E) = -5x^2 - 4x - 2$$

Comparing coefficients,

$$5E = -5 \to E = -1, 4E + 5D = -4 \to D = 0, 2E + 2D + 5C = -2 \to C = 0$$

Hence $y_p = -x^2$ The general solution is then

$$w = w_c + w_p = e^{-x}(A\cos 2x + B\sin 2x) - x^2$$

$$y = \frac{1}{w} \to \frac{y}{f(x)} = \frac{1}{f(x)e^{-x}(A\cos 2x + B\sin 2x) - x^2 f(x)}$$

As $x \to \infty$, $e^{-x} \to 0$, hence $\frac{y}{f(x)} \to -\frac{1}{x^2 f(x)}$

As $x \to \infty, y \to 1$, hence $\lim_{x\to\infty} x^2 f(x) = -1$, a suitable $f(x)$ is:

$$f(x) = -\frac{1}{x^2}$$

CHAPTER 7: COMPLEX NUMBERS

A) Pre-requisite Knowledge: Paper 3 Mathematics

- Converting between 3 forms of complex numbers (cartesian, polar and exponential)
- Basic arithmetic of complex numbers, especially conjugation (involving polar form)
- Know what is an Argand diagram
- Euler's identity: $e^{i\theta} = \cos\theta + i\sin\theta$.

B) de' Moivre's Theorem

- LO: understand de Moivre's theorem, for a positive integral exponent, in terms of the geometrical effect of multiplication of complex numbers;
- LO: prove de Moivre's theorem for a positive integral exponent;

> **If $z = \cos\theta + i\sin\theta$, then $z^n = (\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$.**

- In exponential form: $z^n = (e^{i\theta})^n = e^{in\theta}$.
- For integers *n*, this is the only value for z^n.
- For rational numbers *n*, this is one of the values z^n can take (discussed in part C)
- The proof of this theorem for positive integers uses mathematical induction.

Example: W2006 P1 Q11 (first part: prove de' Moivre's theorem for positive integers)

Base case: $n = 1$ $(\cos\theta + i\sin\theta)^1 = \cos 1\theta + i\sin 1\theta = \cos\theta + i\sin\theta$

Hence the statement is true for $n = 1$.

Assume the statement is true for some integer *k*. $(\cos\theta + i\sin\theta)^k = \cos k\theta + i\sin k\theta$.

To prove: for $n = k + 1$: $(\cos\theta + i\sin\theta)^{k+1} = \cos(k+1)\theta + i\sin(k+1)\theta$

$$\begin{aligned}(\cos\theta + i\sin\theta)^{k+1} &= (\cos\theta + i\sin\theta)^k(\cos\theta + i\sin\theta)\\ &= (\cos k\theta + i\sin k\theta)(\cos\theta + i\sin\theta)\\ &= (\cos k\theta\cos\theta - \sin k\theta\sin\theta) + i\,(\sin k\theta\cos\theta + \cos k\theta\sin\theta)\\ &= \cos(k+1)\theta + i\sin(k+1)\theta\end{aligned}$$

Hence the statement is true for $n = k + 1$. Thus by principle of mathematical induction, $(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$ for all positive integers *n*.

- For negative integers, we show that if it is true for *n*, it is also true for –*n*.

$$(\cos\theta + i\sin\theta)^{-n} = \frac{1}{(\cos\theta + i\sin\theta)^n} = \frac{1}{\cos n\theta + i\sin n\theta} = \frac{\cos n\theta - i\sin n\theta}{cos^2 n\theta + sin^2 n\theta}$$

$$= \cos n\theta - i\sin n\theta = \cos(-n)\theta + i\sin(-n)\theta$$

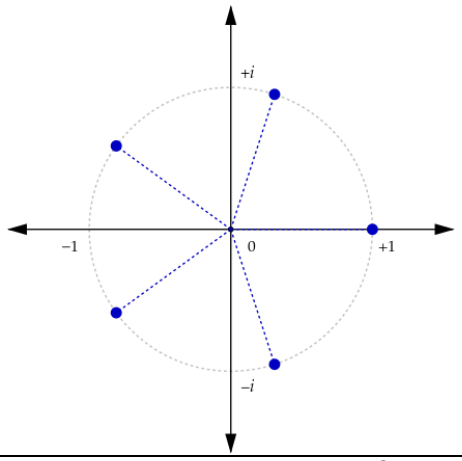
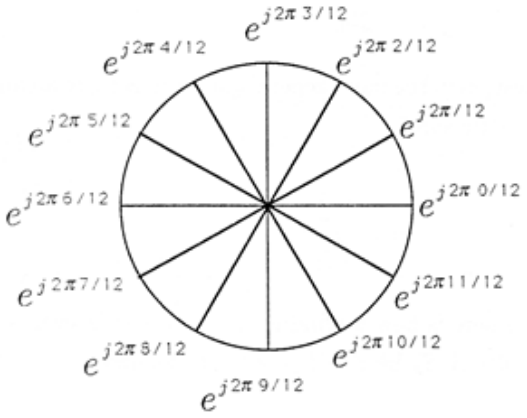
C) Applying de Moivre's Theorem: Powers and Roots

- LO: use de Moivre's theorem, for a positive or negative rational exponent in finding and using the n^{th} roots of unity.

  - To raise a complex number to a high exponent, just change it to polar form and apply de Moivre's theorem.
  - Example: $(1+i)^6 = \left[\sqrt{2}\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right)\right]^6 = 8\left(\cos\frac{6\pi}{4} + i\sin\frac{6\pi}{4}\right) = -8i$

<u>n^{th} Roots of Unity</u>

  - Means the n^{th} roots of 1.
  - For example, cube roots of 1 are $1, \frac{1}{2} + \frac{\sqrt{3}}{2}i$ and $\frac{1}{2} - \frac{\sqrt{3}}{2}i$. The roots can be obtained by solving $x^3 - 1 = 0$.
  - If the n^{th} roots are plotted in an Argand diagram, the lines will form a circle and are $\frac{2\pi}{n}$ away from each other as seen below:

5th roots of unity	**12th roots of unity**
	
Angle between each root is $\frac{2}{5}\pi$	Angle between each root is $\frac{2}{12}\pi$

  - Derivation of the n^{th} roots of unity:

Let $z = 1^{\frac{1}{n}} = r(\cos\theta + i\sin\theta)$.
Raise both sides to the n^{th} power: $z^n = r^n(\cos n\theta + i\sin n\theta) = 1 = \cos 0 + i\sin 0$
Comparing the coefficients: $r^n = 1 \rightarrow r = 1$ since r is real.

$$n\theta = 0, 2\pi, 4\pi, \ldots 2(n-1)\pi \text{ , since we want } n \text{ values of } \theta.$$

$$\theta = \frac{2k\pi}{n}, \text{ where } k = 0,1,2,\ldots,n-1$$

$$z = \cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n} = e^{i\frac{2k\pi}{n}}$$

  - This method can be extended to find roots of other numbers (you can use this to do square root problems in Paper 3 when $n = 2$) as shown in Example 1 below.

Example 1: S2004 P1 Q7

Find the roots of the equation

$$z^3 = -(4\sqrt{3}) + 4\mathrm{i},$$

giving your answers in the form $re^{i\theta}$, where $r > 0$ and $0 \leqslant \theta < 2\pi$. [5]

Denoting these roots by z_1, z_2, z_3, show that, for every positive integer k,

$$z_1^{3k} + z_2^{3k} + z_3^{3k} = 3\left(2^{3k} e^{\frac{5}{6}k\pi i}\right).$$ [4]

First convert the number into polar form.

$$z^3 = -4\sqrt{3} + 4i = r(\cos\theta + i\sin\theta)$$

Comparing coefficients, $r\cos\theta = -4\sqrt{3}, r\sin\theta = 4$. Hence $r = 8$ and $\theta = \frac{5}{6}\pi$.

Let $z = a(\cos\alpha + i\sin\alpha)$. Using de Moivre's theorem,

$$z^3 = a^3(\cos 3\alpha + i\sin 3\alpha) = 8\left(\cos\frac{5}{6}\pi + i\sin\frac{5}{6}\pi\right)$$

Comparing coefficients again, $a^3 = 8 \to a = 2$

$$3\alpha = \frac{5}{6}\pi, \frac{17}{6}\pi, \frac{29}{6}\pi \to \alpha = \frac{5}{18}\pi, \frac{17}{18}\pi, \frac{29}{18}\pi$$

$$z = ae^{i\alpha} = 2e^{i\frac{5}{18}\pi}, 2e^{i\frac{11}{18}\pi}, 2e^{i\frac{17}{18}\pi}$$

Notice that since z_1, z_2 and z_3 are cube roots of $-4\sqrt{3} + 4i$,

$$z_1^3 = z_2^3 = z_3^3 = -4\sqrt{3} + 4i = 2^3 e^{i\frac{5}{6}\pi}$$

$$z_1^{3k} = z_2^{3k} = z_3^{3k} = 2^{3k} e^{i\frac{5k}{6}\pi}$$

$$z_1^{3k} + z_2^{3k} + z_3^{3k} = 3\left(2^{3k} e^{i\frac{5k}{6}\pi}\right)$$

Example 2: W2004 P1 Q6

Write down all the 8th roots of unity. [2]

Verify that

$$(z - e^{i\theta})(z - e^{-i\theta}) \equiv z^2 - (2\cos\theta)z + 1.$$ [1]

Hence express $z^8 - 1$ as the product of two linear factors and three quadratic factors, where all coefficients are real and expressed in a non-trigonometric form. [5]

8th roots of unity: $e^{i\frac{2k\pi}{8}} = e^{\frac{ik\pi}{4}}$, where $k = 0, \pm1, \pm2, \pm3, 4$.

$$(z - e^{i\theta})(z - e^{-i\theta}) = z^2 - (e^{i\theta} + e^{-i\theta})z + 1 = z^2 - (2\cos\theta)z + 1$$

When $z^8 = 1$, z are the 8th roots of unity. Notice that we have

$$\left(z - e^{\frac{i\pi}{4}}\right)\left(z - e^{-\frac{i\pi}{4}}\right) = z^2 - 2\cos\frac{\pi}{4}z + 1 = z^2 - \sqrt{2}z + 1$$

$$\left(z - e^{\frac{i2\pi}{4}}\right)\left(z - e^{-\frac{i2\pi}{4}}\right) = z^2 - 2\cos\frac{2\pi}{4}z + 1 = z^2 + 1$$

$$\left(z - e^{\frac{i3\pi}{4}}\right)\left(z - e^{-\frac{i3\pi}{4}}\right) = z^2 - 2\cos\frac{3\pi}{4}z + 1 = z^2 + \sqrt{2}z + 1$$

[That's why I wrote ±1, ±2, ±3 instead of 1, 2, 3, 5, 6, 7 in part one]. The other 2 factors are

$$\left(z - e^{\frac{i0\pi}{4}}\right) = z - 1 \text{ and } \left(z - e^{\frac{i4\pi}{4}}\right) = z + 1$$

Hence

$$z^8 - 1 \equiv (z+1)(z-1)(z^2 - \sqrt{2}z + 1)(z^2 + \sqrt{2}z + 1)(z^2 + 1)$$

D) Applying de Moivre's Theorem: Trigonometry

- LO: use de Moivre's theorem for positive integral exponent to express trigonometrical ratios of multiple angles in terms of powers of trigonometrical ratios of the fundamental angle;
- LO: use de Moivre's theorem for a positive or negative rational exponent in expressing powers of sin θ and cos θ in terms of multiple angles.

- Involves proving trigonometric identities using de'Moivre's theorem
- Important result: if $z = \cos\theta + i\sin\theta$, we have

$z^n = \cos n\theta + i\sin n\theta$ – (1),

$z^{-n} = \cos(-n)\theta + i\sin(-n)\theta = \cos n\theta - i\sin n\theta$ – (2)

Thus, (1) + (2): $\boldsymbol{z^n + z^{-n} = 2\cos n\theta}$, (1) – (2): $\boldsymbol{z^n - z^{-n} = 2i\sin n\theta}$

Case 1: Converting Powers to Multiples

Strategy: Expand using binomial theorem, group the terms and use the two results above

Main purpose: Integrating trigonometric functions (an alternative to reduction formulae)

Example: W2003 P1 Q8

Given that $z = e^{i\theta}$ and n is a positive integer, show that

$$z^n + \frac{1}{z^n} = 2\cos n\theta \quad \text{and} \quad z^n - \frac{1}{z^n} = 2i\sin n\theta. \qquad [2]$$

Hence express $\sin^6\theta$ in the form

$$p\cos 6\theta + q\cos 4\theta + r\cos 2\theta + s,$$

where the constants p, q, r, s are to be determined. [4]

Hence find the mean value of $\sin^6\theta$ with respect to θ over the interval $0 \leqslant \theta \leqslant \frac{1}{4}\pi$. [5]

First part: See above

Second part: let $n = 1$:

$$(2i\sin\theta)^6 = \left(z - z^{-1}\right)^6$$
$$64i^6\sin^6\theta = z^6 + z^{-6} - 6z^4 - 6z^{-4} + 15z^2 + 15z^{-2} - 20$$
$$-64\sin^6\theta = 2\cos 6\theta - 12\cos 4\theta + 30\cos 2\theta - 20$$
$$\sin^6\theta = -\frac{1}{64}(2\cos 6\theta - 12\cos 4\theta + 15\cos 2\theta - 20)$$
$$= -\frac{1}{32}\cos 6\theta + \frac{3}{16}\cos 4\theta - \frac{15}{32}\cos 2\theta + \frac{5}{16}$$

Third part

$$\text{Width} = \tfrac{1}{4}\pi - 0 = \tfrac{1}{4}\pi$$

$$Area = \int_0^{\frac{1}{4}\pi} \sin^6\theta d\theta$$
$$= \int_0^{\frac{1}{4}\pi} -\frac{1}{32}\cos 6\theta + \frac{3}{16}\cos 4\theta - \frac{15}{32}\cos 2\theta + \frac{5}{16} d\theta$$
$$= \left[-\frac{1}{192}\sin 6\theta + \frac{3}{64}\sin 4\theta - \frac{15}{64}\sin 2\theta + \frac{5}{16}\theta \right]_0^{\frac{1}{4}\pi}$$
$$= \frac{1}{192} - \frac{15}{64} + \frac{5}{64}\pi = \frac{5}{64}\pi - \frac{11}{48}$$

Mean value × width = Area, hence

$$m\left(\frac{1}{4}\pi\right) = \frac{5}{64}\pi - \frac{11}{48} \rightarrow m = \frac{5}{16} - \frac{11}{12\pi}$$

Case 2: Converting Multiples to Powers

General strategy: Expand RHS using Binomial theorem then equate real or imaginary part

Main purpose: Is a hidden yet very useful strategy to solve polynomial equations

Example: W2010 P1 Q10

By using de Moivre's theorem to express $\sin 5\theta$ and $\cos 5\theta$ in terms of $\sin\theta$ and $\cos\theta$, show that

$$\tan 5\theta = \frac{5t - 10t^3 + t^5}{1 - 10t^2 + 5t^4},$$

where $t = \tan\theta$. [5]

Show that the roots of the equation $x^4 - 10x^2 + 5 = 0$ are $\tan\left(\frac{1}{5}n\pi\right)$ for $n = 1, 2, 3, 4$. [2]

By considering the product of the roots of this equation, find the exact value of $\tan\left(\frac{1}{5}\pi\right)\tan\left(\frac{2}{5}\pi\right)$. [3]

$$\cos 5\theta + i\sin 5\theta = (\cos\theta + i\sin\theta)^5$$
$$= \cos^5\theta + 5i\cos^4\theta\sin\theta - 10\cos^3\theta\sin^2\theta - 10i\cos^2\theta\sin^3\theta$$
$$+ 5\cos\theta\sin^4\theta + i\sin^5\theta$$

Equating real coefficients:

$$\cos 5\theta = cos^5\theta - 10cos^3\theta\, sin^2\theta + 5\cos\theta\, sin^4\theta$$

Equating imaginary coefficients:

$$\sin 5\theta = 5cos^4\theta \sin\theta - 10cos^2\theta\, sin^3\theta + sin^5\theta$$

Thus $\tan 5\theta = \frac{\sin 5\theta}{\cos 5\theta} = \frac{5\cos^4\theta\sin\theta - 10\cos^2\theta\sin^3\theta + \sin^5\theta}{\cos^5\theta - 10\cos^3\theta\sin^2\theta + 5\cos\theta\sin^4\theta}$

Dividing by $cos^5\theta$,

$$\tan 5\theta = \frac{5\frac{\sin\theta}{\cos\theta} - 10\frac{sin^3\theta}{cos^3\theta} + \frac{sin^5\theta}{cos^5\theta}}{1 - 10\frac{sin^2\theta}{cos^2\theta} + 5\frac{sin^4\theta}{cos^4\theta}} = \frac{5\tan\theta - 10\, tan^3\theta + tan^5\theta}{1 - 10tan^2\theta + 5tan^4\theta}$$

Let $t = \tan\theta$: $\tan 5\theta = \frac{5t - 10t^3 + t^5}{1 - 10t^2 + 5t^4}$.

When $\tan 5\theta = 0$, $5\theta = n\pi$ for $n = 0,1,2,3,4$.

Hence $\theta = \frac{1}{5}n\pi, n = 0,1,2,3,4$.

$$\tan 5\theta = 0 \rightarrow 5t - 10t^3 + t^5 = 0$$
$$t(t^4 - 10t^2 + 5) = 0$$

When $t \neq 0, \theta \neq 0$, hence roots of $t^4 - 10t^2 + 5 = 0$ are $t = \tan\frac{1}{5}n\pi$, $n = 1,2,3,4$.

Let $x = t \rightarrow$ hence roots of $x^4 - 10x^2 + 5 = 0$ are $x = \tan\frac{1}{5}n\pi$, $n = 1,2,3,4$.

First notice that $\tan x + \tan(2\pi - x) = \tan 2\pi\,(1 - \tan x \tan(2\pi - x)) = 0$.

So we have $\tan x = -\tan(2\pi - x)$.

Consider product of roots:

$$\tan\frac{1}{5}\pi \tan\frac{2}{5}\pi \tan\frac{3}{5}\pi \tan\frac{4}{5}\pi = 5$$

$$\tan\frac{1}{5}\pi \tan\frac{2}{5}\pi \left(-\tan\frac{2}{5}\pi\right)\left(-\tan\frac{1}{5}\pi\right) = 5$$

$$\left(\tan\frac{1}{5}\pi\right)^2 \left(\tan\frac{2}{5}\pi\right)^2 = 5 \rightarrow \tan\frac{1}{5}\pi \tan\frac{2}{5}\pi = \sqrt{5}$$

, since both $\tan\frac{1}{5}\pi$ and $\tan\frac{2}{5}\pi$ are positive.

E) Applying de Moivre's Theorem: Summation of Series

- LO: use de Moivre's theorem, for a positive or negative rational exponent in the summation of series.
  - Main strategy: expand one side using de'Moivre's theorem
  - Then expand the other side using geometric series or binomial theorem
  - Then equate the real part (cosine) or imaginary part (sine) for the complete proof
  - You can also differentiate / integrate both sides of the equation or substitute special values to prove specific cases.

Example 1: S2009 P1 Q12 EITHER

By considering $\sum_{k=0}^{n-1}(1 + \mathrm{i}\tan\theta)^k$, show that

$$\sum_{k=0}^{n-1} \cos k\theta \sec^k\theta = \cot\theta \sin n\theta \sec^n\theta,$$

provided θ is not an integer multiple of $\frac{1}{2}\pi$. [7]

Hence or otherwise show that

$$\sum_{k=0}^{n-1} 2^k \cos\left(\tfrac{1}{3}k\pi\right) = \frac{2^n}{\sqrt{3}} \sin\left(\tfrac{1}{3}n\pi\right).$$ [2]

Given that $0 < x < 1$, show that

$$\sum_{k=0}^{n-1} \frac{\cos(k\cos^{-1}x)}{x^k} = \frac{\sin(n\cos^{-1}x)}{x^{n-1}\sqrt{(1-x^2)}}.$$ [4]

First convert the given term to use de Moivre's theorem.

$$1 + i\tan\theta = \sec\theta\,(\cos\theta + i\sin\theta)$$

$$\therefore \sum_{k=0}^{n-1}(1 + i\tan\theta)^k = \sum_{k=0}^{n-1}[\sec\theta\,(\cos\theta + i\sin\theta)]^k = \sum_{k=0}^{n-1} sec^k\,\theta\,(\cos k\theta + i\sin k\theta)$$

Take the real part for LHS:

$$Re\left[\sum_{k=0}^{n-1}(1+i\tan\theta)^k\right]=\sum_{k=0}^{n-1}\cos k\theta\, sec^k\theta$$

Then we find RHS using geometric series:

$$\sum_{k=0}^{n-1}(1+i\tan\theta)^k=\frac{(1+i\tan\theta)^n-1}{(1+i\tan\theta)-1}\cdot\frac{-i}{-i}=\frac{[\sec^n\theta(\cos n\theta+i\sin n\theta)-1](-i)}{\tan\theta}$$

Since the denominator is real, we just pick terms that make the real part for the numerator for the final answer.

$$Re\left[\sum_{k=0}^{n-1}(1+i\tan\theta)^k\right]=\frac{sec^n\theta\sin n\theta}{\tan\theta}=\cot\theta\sin n\theta\, sec^n\theta$$

Hence LHS = RHS:

$$Re\left[\sum_{k=0}^{n-1}(1+i\tan\theta)^k\right]=\sum_{k=0}^{n-1}\cos k\theta\, sec^k\theta=\cot\theta\sin n\theta\, sec^n\theta$$

Let $\theta=\frac{1}{3}\pi$:

$$\sum_{k=0}^{n-1}\cos\frac{1}{3}k\pi\left(\sec\frac{1}{3}\pi\right)^k=\cot\frac{1}{3}\pi\sin\frac{1}{3}n\pi\left(\sec\frac{1}{3}\pi\right)^n$$

$$\sum_{k=0}^{n-1}2^k\cos\frac{1}{3}k\pi=\frac{2^n}{\sqrt{3}}\sin\frac{1}{3}n\pi$$

Let $\theta=\cos^{-1}x$: First we evaluate a few things:

$$\theta=\cos^{-1}x\to\cos\theta=x\to cos^k\theta=x^k\to sec^k\theta=\frac{1}{x^k}$$

$$\sin\theta=\sqrt{1-x^2}\to\cot\theta=\frac{\cos\theta}{\sin\theta}=\frac{x}{\sqrt{1-x^2}}$$

Then substitute.

$$\sum_{k=0}^{n-1}\cos k(\cos^{-1}x)\left(\frac{1}{x^k}\right)=\frac{x}{\sqrt{1-x^2}}\sin(n\cos^{-1}x)\left(\frac{1}{x^n}\right)$$

$$\sum_{k=0}^{n-1}\left(\frac{\cos k(\cos^{-1}x)}{x^k}\right)=\frac{\sin(n\cos^{-1}x)}{x^{n-1}\sqrt{1-x^2}}$$

Example 2: S2008 P1 Q10

By considering $\sum_{n=1}^{N}z^{2n-1}$, where $z=e^{i\theta}$, show that

$$\sum_{n=1}^{N}\cos(2n-1)\theta=\frac{\sin(2N\theta)}{2\sin\theta},$$

where $\sin\theta\neq0$. [6]

Deduce that

$$\sum_{n=1}^{N}(2n-1)\sin\left[\frac{(2n-1)\pi}{N}\right]=-N\operatorname{cosec}\frac{\pi}{N}.$$ [4]

LHS: Use de' Moivre's theorem, then take the real part

$$\sum_{n=1}^{N}z^{2n-1}=\sum_{n=1}^{N}(\cos\theta+i\sin\theta)^{2n-1}=\sum_{n=1}^{N}\cos(2n-1)\theta+i\sin(2n-1)\theta$$

$$Re\left(\sum_{n=1}^{N} z^{2n-1}\right) = \sum_{n=1}^{N} \cos(2n-1)\theta$$

RHS: Use geometric series

$$\sum_{n=1}^{N} z^{2n-1} = \sum_{n=1}^{N} (e^{i\theta})^{2n-1} = \frac{e^{i\theta}(e^{2i\theta^N} - 1)}{e^{2i\theta} - 1} = \frac{e^{i2N\theta} - 1}{e^{i\theta} - e^{-i\theta}} = \frac{\cos 2N\theta + i \sin 2N\theta - 1}{2i \sin\theta}$$

Take the real part:

$$Re\left(\sum_{n=1}^{N} z^{2n-1}\right) = \frac{i \sin 2N\theta}{2i \sin\theta} = \frac{\sin 2N\theta}{2 \sin\theta}$$

$$Re\left(\sum_{n=1}^{N} z^{2n-1}\right) = \sum_{n=1}^{N} \cos(2n-1)\theta = \frac{\sin 2N\theta}{2 \sin\theta}$$

Differentiate both sides with respect to θ:

$$\sum_{n=1}^{N} -(2n-1)\sin(2n-1)\theta = \frac{\sin\theta\,(2N\cos 2N\theta) - \cos\theta\,(\sin 2N\theta)}{2sin^2\theta}$$

Let $\theta = \frac{\pi}{N}$:

$$-\sum_{n=1}^{N} (2n-1)\sin\frac{(2n-1)\pi}{N} = \frac{\sin\frac{\pi}{N}(2N\cos 2\pi) - \cos\frac{\pi}{N}(\sin 2\pi)}{2sin^2\frac{\pi}{N}} = \frac{2N}{2\sin\frac{\pi}{N}}$$

$$\sum_{n=1}^{N} (2n-1)\sin\frac{(2n-1)\pi}{N} = -\frac{N}{\sin\frac{\pi}{N}} = -N\csc\frac{\pi}{N}$$

Example 3: W2012 P13 Q8

Let $z = \cos\theta + \mathrm{i}\sin\theta$. Show that

$$1 + z = 2\cos\tfrac{1}{2}\theta(\cos\tfrac{1}{2}\theta + \mathrm{i}\sin\tfrac{1}{2}\theta). \qquad [3]$$

By considering $(1 + z)^n$, where n is a positive integer, deduce the sum of the series

$$\binom{n}{1}\sin\theta + \binom{n}{2}\sin 2\theta + \ldots + \binom{n}{n}\sin n\theta. \qquad [6]$$

$$1+z = 1+\cos\theta+i\sin\theta = 1+2\cos^2\frac{1}{2}\theta - 1 + 2i\cos\frac{1}{2}\theta\sin\frac{1}{2}\theta = 2\cos\frac{1}{2}\theta\left(\cos\frac{1}{2}\theta + i\sin\frac{1}{2}\theta\right)$$

LHS: Binomial theorem: $(1+z)^n = \binom{n}{0}1^n z^0 + \binom{n}{1}1^{n-1}z^1 + \binom{n}{2}1^{n-2}z^2 \ldots + \binom{n}{n}1^0 z^n$

$Im(z^k) = \sin k\theta$, $1^k = 1$ for all integers k, hence

$$Im(1+z)^n = 1 + \binom{n}{1}\sin\theta + + \binom{n}{2}\sin 2\theta + \cdots + \binom{n}{n}\sin n\theta$$

RHS: de Moivre's theorem

$$(1+z)^n = \left[2\cos\frac{1}{2}\theta\left(\cos\frac{1}{2}\theta + i\sin\frac{1}{2}\theta\right)\right]^n = 2^n cos^n\frac{1}{2}\theta\left(cos\ \frac{1}{2}n\theta + i\sin\frac{1}{2}n\theta\right)$$

Taking imaginary parts:

$$Im(1+z)^n = 2^n cos^n\frac{1}{2}\theta\sin\frac{1}{2}n\theta$$

$$1 + \binom{n}{1}\sin\theta + \binom{n}{2}\sin 2\theta + \cdots + \binom{n}{n}\sin n\theta = 2^n cos^n\frac{1}{2}\theta\sin\frac{1}{2}n\theta$$

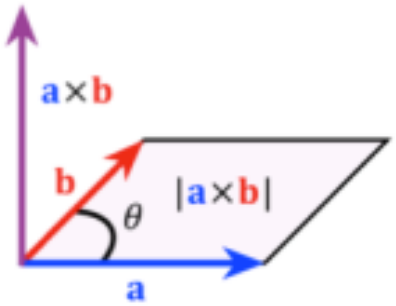
$$\binom{n}{1}\sin\theta + \binom{n}{2}\sin 2\theta + \cdots + \binom{n}{n}\sin n\theta = 2^n cos^n\frac{1}{2}\theta\sin\frac{1}{2}n\theta - 1$$

CHAPTER 8: VECTORS

8.1 New Vector Terminologies

A) The Cross Product / The Vector Product

- LO: recall that the vector product $\boldsymbol{a} \times \boldsymbol{b}$ of two vectors can be expressed either as $|\mathbf{a}||\mathbf{b}| \sin\theta\, \hat{\boldsymbol{n}}$, where $\hat{\boldsymbol{n}}$ is a unit vector, or in component form as $(a_2b_3 - a_3b_2)\mathbf{i} + (a_3b_1 - a_1b_3)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k}$.



- The vector product of two vectors, $\boldsymbol{a}$ and $\boldsymbol{b}$ is denoted by $\boldsymbol{a} \times \boldsymbol{b}$ or sometimes as $\boldsymbol{a}$ ^ $\boldsymbol{b}$.
- It produces a vector that is perpendicular to both vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, while the dot product only produces a scalar / constant.
- By definition, $\boldsymbol{a} \times \boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}| \sin\theta\, \hat{\boldsymbol{n}}$, where θ is the angle between the two vectors and $\hat{\boldsymbol{n}}$ is a unit vector perpendicular to both $\boldsymbol{a}$ and $\boldsymbol{b}$.
- The direction of $\boldsymbol{a} \times \boldsymbol{b}$ can be determined using your right hand: if your thumb is pointing towards vector $\boldsymbol{a}$ and index finger is pointing towards vector $\boldsymbol{b}$, the middle finger will point towards vector $\boldsymbol{a} \times \boldsymbol{b}$.
- Rules of the Vector Product:
  1. $\boldsymbol{a} \times \boldsymbol{b} = -\boldsymbol{b} \times \boldsymbol{a}$ (cross product is anti-commutative)
  2. $\boldsymbol{a} \times (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \times \boldsymbol{b} + \boldsymbol{a} \times \boldsymbol{c}$ (cross product is distributive over addition)
  3. $s(\boldsymbol{a} \times \boldsymbol{b}) = (s\boldsymbol{a}) \times \boldsymbol{b} + \boldsymbol{a} \times (s\boldsymbol{b})$, for any constant s.
  4. $\boldsymbol{i} \times \boldsymbol{i} = \boldsymbol{j} \times \boldsymbol{j} = \boldsymbol{k} \times \boldsymbol{k} = 0$ ($\theta = 0$ as these vectors are parallel)
  5. $\boldsymbol{i} \times \boldsymbol{j} = \boldsymbol{k},\ \boldsymbol{j} \times \boldsymbol{k} = \boldsymbol{i},\ \boldsymbol{k} \times \boldsymbol{i} = \boldsymbol{j}$ (by applying the right-hand rule)
  6. $\boldsymbol{j} \times \boldsymbol{i} = -\boldsymbol{k}, \boldsymbol{k} \times \boldsymbol{j} = -\boldsymbol{i}, \boldsymbol{i} \times \boldsymbol{k} = -\boldsymbol{j}$ (Rule 1 + Rule 5)
  7. $\boldsymbol{a}.\boldsymbol{b} \times \boldsymbol{c} = \boldsymbol{a}.(\boldsymbol{b} \times \boldsymbol{c})$ (cross product is always performed before dot product)

Expressing the vector product in Component form:

- Let $a = a_1\boldsymbol{i} + a_2\boldsymbol{j} + a_3\boldsymbol{k}$ and $\boldsymbol{b} = b_1\boldsymbol{i} + b_2\boldsymbol{j} + b_3\boldsymbol{k}$,

$$
\begin{aligned}
\boldsymbol{a} \times \boldsymbol{b} &= (a_1\boldsymbol{i} + a_2\boldsymbol{j} + a_3\boldsymbol{k}) \times (b_1\boldsymbol{i} + b_2\boldsymbol{j} + b_3\boldsymbol{k}) \\
&= (a_1\boldsymbol{i} + a_2\boldsymbol{j} + a_3\boldsymbol{k}) \times b_1\boldsymbol{i} + (a_1\boldsymbol{i} + a_2\boldsymbol{j} + a_3\boldsymbol{k}) \times b_2\boldsymbol{j} \qquad \text{Rule 2} \\
&\quad + (a_1\boldsymbol{i} + a_2\boldsymbol{j} + a_3\boldsymbol{k}) \times b_3\boldsymbol{k} \\
&= a_1\boldsymbol{i} \times b_1\boldsymbol{i} + a_2\boldsymbol{j} \times b_1\boldsymbol{i} + a_3\boldsymbol{k} \times b_1\boldsymbol{i} + a_1\boldsymbol{i} \times b_2\boldsymbol{j} + a_2\boldsymbol{j} \times b_2\boldsymbol{j} \qquad \text{Rule 2} \\
&\quad + a_3\boldsymbol{k} \times b_2\boldsymbol{j} + a_1\boldsymbol{i} \times b_3\boldsymbol{k} + a_2\boldsymbol{j} \times b_3\boldsymbol{k} + a_3\boldsymbol{k} \times b_3\boldsymbol{k}
\end{aligned}
$$

$$= a_1b_1(\boldsymbol{i}\times\boldsymbol{i}) + a_2b_1(\boldsymbol{j}\times\boldsymbol{i}) + a_3b_1(\boldsymbol{k}\times\boldsymbol{i}) \qquad \text{Rule 3}$$
$$+a_1b_2(\boldsymbol{i}\times\boldsymbol{j}) + a_2b_2(\boldsymbol{j}\times\boldsymbol{j}) + a_3b_2(\boldsymbol{k}\times\boldsymbol{j})$$
$$+a_1b_3(\boldsymbol{i}\times\boldsymbol{k}) + a_2b_3(\boldsymbol{j}\times\boldsymbol{k}) + a_3b_3(\boldsymbol{k}\times\boldsymbol{k})$$
$$= a_1b_1(\boldsymbol{0}) + a_2b_1(-\boldsymbol{k}) + a_3b_1(\boldsymbol{j}) + a_1b_2(\boldsymbol{k}) + a_2b_2(\boldsymbol{0}) \qquad \text{Rule 4,}$$
$$+a_3b_2(-\boldsymbol{i}) + a_1b_3(-\boldsymbol{j}) + a_2b_3(\boldsymbol{i}) + a_3b_3(\boldsymbol{0}) \qquad \text{Rule 5,}$$
$$= (a_2b_3 - a_3b_2)\boldsymbol{i} + (a_3b_1 - a_1b_3)\boldsymbol{j} + (a_1b_2 - a_2b_1)\boldsymbol{k} \qquad \text{Rule 6}$$

- This can also be stated in other forms, like:

Column vector form	$\begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix}\times\begin{pmatrix}b_1\\b_2\\b_3\end{pmatrix} = \begin{pmatrix}a_2b_3 - a_3b_2\\a_3b_1 - a_1b_3\\a_1b_2 - a_2b_1\end{pmatrix} = \begin{pmatrix}a_2b_3 - a_3b_2\\-(a_1b_3 - a_3b_1)\\a_1b_2 - a_2b_1\end{pmatrix}$
Determinant form	$(a_1\boldsymbol{i} + a_2\boldsymbol{j} + a_3\boldsymbol{k})\times(b_1\boldsymbol{i} + b_2\boldsymbol{j} + b_3\boldsymbol{k})$ $= \begin{vmatrix}\boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k}\\a_1 & a_2 & a_3\\b_1 & b_2 & b_3\end{vmatrix} = \boldsymbol{i}\begin{vmatrix}a_2 & a_3\\b_2 & b_3\end{vmatrix} - \boldsymbol{j}\begin{vmatrix}a_1 & a_3\\b_1 & b_3\end{vmatrix} + \boldsymbol{k}\begin{vmatrix}a_1 & a_2\\b_1 & b_2\end{vmatrix}$

- Example: $(3\boldsymbol{i} + 4\boldsymbol{j} + 2\boldsymbol{k})\times(3\boldsymbol{i} - \boldsymbol{j} + \boldsymbol{k})$

$$(3\boldsymbol{i} + 4\boldsymbol{j} + 2\boldsymbol{k})\times(3\boldsymbol{i} - \boldsymbol{j} + \boldsymbol{k}) = \boldsymbol{i}\begin{vmatrix}4 & 2\\-1 & 1\end{vmatrix} - \boldsymbol{j}\begin{vmatrix}3 & 2\\3 & 1\end{vmatrix} + \boldsymbol{k}\begin{vmatrix}3 & 4\\3 & -1\end{vmatrix} = 6\boldsymbol{i} + 3\boldsymbol{j} - 15\boldsymbol{k}$$

$$\begin{pmatrix}3\\4\\2\end{pmatrix}\times\begin{pmatrix}3\\-1\\1\end{pmatrix} = \begin{pmatrix}4(1) - 2(-1)\\-(3(1) - 2(3))\\3(-1) - 4(3)\end{pmatrix} = \begin{pmatrix}6\\3\\-15\end{pmatrix}$$

Application of the Vector Product: Area of a Triangle / Parallelogram

- Since $\boldsymbol{a}\times\boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}|\sin\theta\,\hat{\boldsymbol{n}}, |\boldsymbol{a}\times\boldsymbol{b}| = |\boldsymbol{a}||\boldsymbol{b}|\sin\theta$ as $|\hat{\boldsymbol{n}}| = 1$.
- Area of a triangle $= \frac{1}{2}|\boldsymbol{a}||\boldsymbol{b}|\sin\theta = \frac{1}{2}|\boldsymbol{a}\times\boldsymbol{b}|$.
- Area of a parallelogram = base × perpendicular height $= |\boldsymbol{a}||\boldsymbol{b}|\sin\theta = |\boldsymbol{a}\times\boldsymbol{b}|$.

Example: W2011 P11 Q2

The position vectors of points A, B, C, relative to the origin O, are $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, where

$$\mathbf{a} = 3\mathbf{i} + 2\mathbf{j} - \mathbf{k}, \qquad \mathbf{b} = 4\mathbf{i} - 3\mathbf{j} + 2\mathbf{k}, \qquad \mathbf{c} = 3\mathbf{i} - \mathbf{j} - \mathbf{k}.$$

Find $\mathbf{a}\times\mathbf{b}$ and deduce the area of the triangle OAB. [3]

Hence find the volume of the tetrahedron $OABC$, given that the volume of a tetrahedron is $\frac{1}{3}\times$ area of base $\times$ perpendicular height. [2]

$$(3\boldsymbol{i} + 2\boldsymbol{j} - \boldsymbol{k})\times(4\boldsymbol{i} - 3\boldsymbol{j} + 2\boldsymbol{k}) = \boldsymbol{i}\begin{vmatrix}2 & -1\\-3 & 2\end{vmatrix} - \boldsymbol{j}\begin{vmatrix}3 & -1\\4 & 2\end{vmatrix} + \boldsymbol{k}\begin{vmatrix}3 & 2\\4 & -3\end{vmatrix} = \boldsymbol{i} - 10\boldsymbol{j} - 17\boldsymbol{k}$$

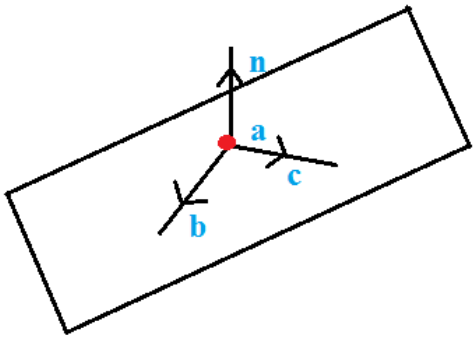
$$Area = \frac{1}{2}\sqrt{1^2 + 10^2 + 17^2} = \frac{1}{2}\sqrt{390}$$

Perpendicular distance from C to plane OAB: $\frac{|3(1) - 1(-10) - 1(-17)|}{\sqrt{1^2 + 10^2 + 17^2}} = \frac{30}{\sqrt{390}}$

$$Volume = \frac{1}{3}\left(\frac{1}{2}\sqrt{390}\right)\left(\frac{30}{\sqrt{390}}\right) = 5$$

B) Equation of a Plane

- LO: understand the significance of all the symbols used when the equation of a plane is expressed in any of the forms $ax + by + cz = d$ or $\mathbf{r} \cdot \mathbf{n} = \mathbf{p}$ or $\mathbf{r} = \mathbf{a} + s\mathbf{b} + t\mathbf{c}$, and convert equations of planes from one form to another in solving problems;



- Vector equation of a plane: $\boldsymbol{r} = \boldsymbol{a} + s\boldsymbol{b} + t\boldsymbol{c}$, where:

$\boldsymbol{r}$ is any point in the plane;
$\boldsymbol{a}$ is a known point in the plane;
$\boldsymbol{b}$ and $\boldsymbol{c}$ are any two vectors in the plane;
$\boldsymbol{s}$ and $\boldsymbol{t}$ are parameters / free variables.

- Notice the difference between a plane equation and a line equation; a plane equation has two parameters while a line equation only has one.
- If we introduce the normal vector of the plane $\boldsymbol{n}$ into the plane equation ($\boldsymbol{n} = \boldsymbol{a} \times \boldsymbol{b}$ as it is perpendicular to all vectors in the plane)

$$\boldsymbol{r}.\boldsymbol{n} = (\boldsymbol{a} + s\boldsymbol{b} + t\boldsymbol{c}).\boldsymbol{n} = \boldsymbol{a}.\boldsymbol{n} + s(\boldsymbol{b}.\boldsymbol{n}) + t(\boldsymbol{c}.\boldsymbol{n})$$

Since $\boldsymbol{n}$ is perpendicular to all vectors in the plane, $\boldsymbol{b}.\boldsymbol{n} = \boldsymbol{c}.\boldsymbol{n} = 0$, so we have

$$\boldsymbol{r}.\boldsymbol{n} = \boldsymbol{a}.\boldsymbol{n}$$

- If we let $\boldsymbol{a}.\boldsymbol{n}$ to be a constant d, let $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$ and let $\boldsymbol{n} = a\boldsymbol{i} + b\boldsymbol{j} + c\boldsymbol{k}$;

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}.\begin{pmatrix} a \\ b \\ c \end{pmatrix} = d \rightarrow ax + by + cz = d$$

- You should be familiar with all three forms of the equation of the plane, and the conversion between each forms of the equation.

Example (W2012 P13 Q9 first part): Convert the plane equation $\boldsymbol{r} = 2\boldsymbol{i} - 3\boldsymbol{j} + \boldsymbol{k} + \lambda(\boldsymbol{i} - 2\boldsymbol{j} - \boldsymbol{k}) + \mu(\boldsymbol{i} + 2\boldsymbol{j} - 2\boldsymbol{k})$ to cartesian form.

$$\boldsymbol{n} = (\boldsymbol{i} - 2\boldsymbol{j} - \boldsymbol{k}) \times (\boldsymbol{i} + 2\boldsymbol{j} - 2\boldsymbol{k}) = \begin{vmatrix} -2 & -1 \\ 2 & -2 \end{vmatrix}\boldsymbol{i} - \begin{vmatrix} 1 & -1 \\ 1 & -2 \end{vmatrix}\boldsymbol{j} + \begin{vmatrix} 1 & -2 \\ 1 & 2 \end{vmatrix}\boldsymbol{k} = 6\boldsymbol{i} + \boldsymbol{j} + 4\boldsymbol{k}$$

$$\boldsymbol{a}.\boldsymbol{n} = (2\boldsymbol{i} - 3\boldsymbol{j} + \boldsymbol{k}).(6\boldsymbol{i} + \boldsymbol{j} + 4\boldsymbol{k}) = 12 - 3 + 4 = 13$$

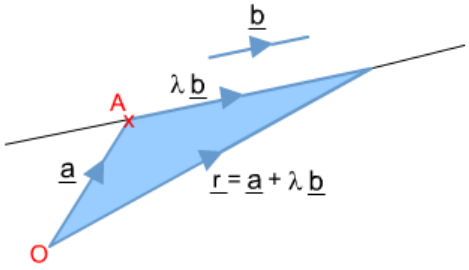
The equation is $\boldsymbol{r}.(6\boldsymbol{i} + \boldsymbol{j} + 4\boldsymbol{k}) = 13 \rightarrow 6x + y + 4z = 13$

8.2 Vector Geometry

- LO: use equations of lines and planes, together with scalar and vector products where appropriate, to solve problems concerning distances, angles and intersections, including:
  - determining whether a line lies in a plane, is parallel to a plane or intersects a plane, and finding the point of intersection of a line and a plane when it exists,
  - finding the perpendicular distance from a point to a plane, and from a point to a line,
  - finding the angle between a line and a plane, and the angle between two planes,
  - finding an equation for the line of intersection of two planes.
  - calculating the shortest distance between two skew lines,
  - finding an equation for the common perpendicular to two skew lines.
  - ➢ In this section we will review some techniques used in solving P3 Vectors problems (as most of its contents overlap with FM Vectors)

A) Equations

1. Equation of a line



- ➢ The vector equation of a line is $\boldsymbol{r} = \boldsymbol{a} + t\boldsymbol{b}$ where:
  $\boldsymbol{a}$ is a fixed point of the line (a position vector)
  $\boldsymbol{b}$ is any vector parallel to the line (a direction vector)
  t is a parameter / free variable
- ➢ The Cartesian equation of the line can be obtained by equalling the parameters.

Let $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$, $\boldsymbol{a} = x_0\boldsymbol{i} + y_0\boldsymbol{j} + z_0\boldsymbol{k}$, $\boldsymbol{b} = l\boldsymbol{i} + m\boldsymbol{j} + n\boldsymbol{k}$. Thus

$$x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k} = x_0\boldsymbol{i} + y_0\boldsymbol{j} + z_0\boldsymbol{k} + t(l\boldsymbol{i} + m\boldsymbol{j} + n\boldsymbol{k})$$

Equating each component:

$$x = x_0 + tl \qquad y = y_0 + tm \qquad z = z_0 + tn$$

$$t = \frac{x - x_0}{l} - (1) \qquad t = \frac{y - y_0}{m} - (2) \qquad t = \frac{z - z_0}{n} - (3)$$

$(1) = (2) = (3)$:

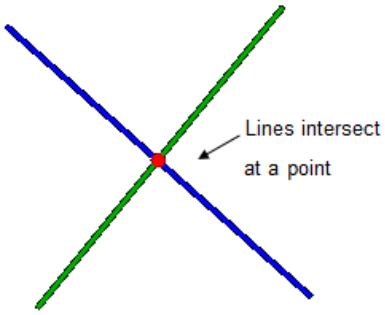
$$\boxed{\frac{\boldsymbol{x - x_0}}{\boldsymbol{l}} = \frac{\boldsymbol{y - y_0}}{\boldsymbol{m}} = \frac{\boldsymbol{z - z_0}}{\boldsymbol{n}}}$$

2. Equation of a plane (discussed in the previous section)

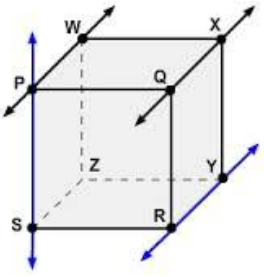
B) Intersections

Case 1: Intersection of two lines

- ➢ Two lines intersect at a **point**. The point can be found by solving both line equations simultaneously.
- ➢ If they do not intersect, they are either parallel or skew.



Two lines are parallel if they have the same direction vector.



In the diagram above, the lines PS and RY are skew.

Example: Given the equations of three lines below:

$$A\text{: } \boldsymbol{r} = 2\boldsymbol{i} + \boldsymbol{j} + 4\boldsymbol{k} + s(\boldsymbol{i} + \boldsymbol{j} - \boldsymbol{k})$$

$$B\text{: } \boldsymbol{r} = -2\boldsymbol{i} + 2\boldsymbol{j} + \boldsymbol{k} + t(-2\boldsymbol{i} + \boldsymbol{j} + \boldsymbol{k})$$

$$C\text{: } \boldsymbol{r} = 6\boldsymbol{i} + 2\boldsymbol{j} + \boldsymbol{k} + p(-2\boldsymbol{i} + \boldsymbol{j} + \boldsymbol{k})$$

1. Lines A and C intersect (verify that the point of intserection is $4\boldsymbol{i} + 3\boldsymbol{j} + 2\boldsymbol{k}$)

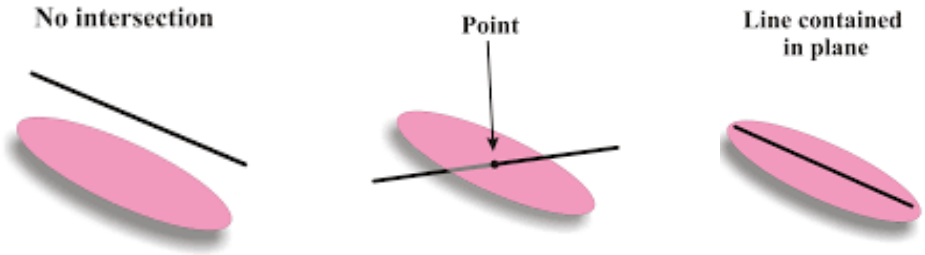
2. Lines B and C are parallel (same direction vector but different position vectors)

3. Lines A and B are skew,

Equating coefficients gives: $1 + s = 2 + t,\ 4 - s = 1 + t$. Solving gives $t = 1, s = 2$.
The solution does not satisfy the first set of equations: $2 + 2 \neq -2 - 2(1)$,
thus no intersection points. The direction vectors are different, hence the lines are skew.

Case 2: Intersection of a line and a plane

- A line and a plane will intersect at a **point**. The point can be found by solving both equations simultaneously.
- If the line and the plane do not intersect, then the line is either parallel to or coincides with the plane.



Example 2: Plane P has equation $6x + y - 8z = 6$. Lines D, E and F have equations

$$D\text{: } \boldsymbol{r} = \boldsymbol{i} + 5\boldsymbol{j} - 4\boldsymbol{k} + s(\boldsymbol{i} - 3\boldsymbol{j} + 5\boldsymbol{k})$$

$$E\text{: } \boldsymbol{r} = -2\boldsymbol{i} + 2\boldsymbol{j} + \boldsymbol{k} + t(\boldsymbol{i} + 2\boldsymbol{j} + \boldsymbol{k})$$

$$F\text{: } \boldsymbol{r} = 4\boldsymbol{i} - 2\boldsymbol{j} + 2\boldsymbol{k} + p(\boldsymbol{i} + 2\boldsymbol{j} + \boldsymbol{k})$$

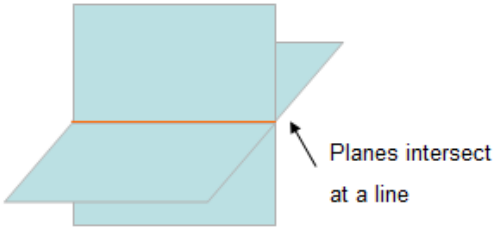
1. Line D intersects with plane P (verify that the point of intserection is $2\boldsymbol{i} + 2\boldsymbol{j} + \boldsymbol{k}$)

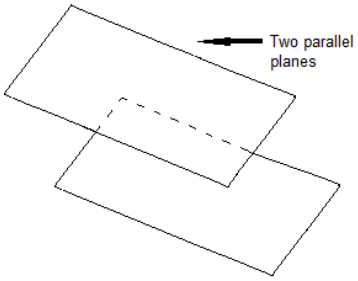
2. Line E and plane P are parallel $[6(-2 + t) + (2 + 2t) - 8(1 + t) = -18 \neq 6]$

3. Line F and plane P coincide $[6(4 + p) + (-2 + 2p) - 8(2 + p) = 6]$

Case 3: Intersection of two planes

- Two planes intersect at a **line** if they are not parallel.





Example: Given the equations of three planes below:

$\pi_1: 6x + y + 4z = 13 \qquad \pi_2: 6x + y + 4z = 25 \qquad \pi_3: 3x - 2y - 3z = 4$

1. Planes π_1 and π_2 are parallel, as they both have normal vector $6\boldsymbol{i} + \boldsymbol{j} + 4\boldsymbol{k}$ and different constants.

2. Find an equation for the line of intersection of planes π_1 and π_3.

Method 1: P3 (Solving for 2 points)	Method 2: FM (1 point + cross product)
$(1) - 2(3): 5y + 10z = 5$ Two possible solutions are: $y = 1, z = 0$ and $y = -1, z = 1$. Substituting gives $x = 2$ and $x = \frac{5}{3}$. Direction vector $= \begin{pmatrix}2\\1\\0\end{pmatrix} - \begin{pmatrix}5/3\\-1\\1\end{pmatrix} = \begin{pmatrix}1/3\\2\\-1\end{pmatrix}$. Equation is $\boldsymbol{r} = 2\boldsymbol{i} + \boldsymbol{j} + p(\frac{1}{3}\boldsymbol{i} + 2\boldsymbol{j} - \boldsymbol{k})$.	A point on both the line and the plane is $(2, 1, 0)$ from Method 1. $(6\boldsymbol{i} + \boldsymbol{j} + 4\boldsymbol{k}) \times (3\boldsymbol{i} - 2\boldsymbol{j} - 3\boldsymbol{k})$ $= \begin{vmatrix}1 & 4\\-2 & -3\end{vmatrix}\boldsymbol{i} - \begin{vmatrix}6 & 4\\3 & -3\end{vmatrix}\boldsymbol{j} + \begin{vmatrix}6 & 1\\3 & -2\end{vmatrix}\boldsymbol{k}$ $= 5\boldsymbol{i} + 30\boldsymbol{j} - 15\boldsymbol{k}$ $= 5(\boldsymbol{i} + 6\boldsymbol{j} - 3\boldsymbol{k})$ The line is $\boldsymbol{r} = 2\boldsymbol{i} + \boldsymbol{j} + t(\boldsymbol{i} + 6\boldsymbol{j} - 3\boldsymbol{k})$.

Method 3: FM Method

$$(1) - 2(3): 5y + 10z = 5$$

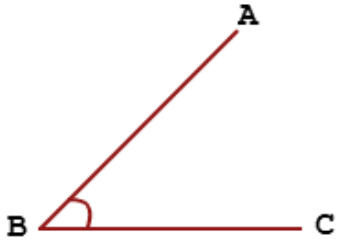
Let $z = -3t$, where $t \in R$. Thus $= \frac{5+30t}{5} = 1 + 6t$, $x = \frac{1}{6}(13 - 1 - 6t + 12t) = 2 + t$.

Thus the line is $\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}2+t\\1+6t\\-3t\end{pmatrix} = \begin{pmatrix}2\\1\\0\end{pmatrix} + t\begin{pmatrix}1\\6\\-3\end{pmatrix} \rightarrow \boldsymbol{r} = 2\boldsymbol{i} + \boldsymbol{j} + t(\boldsymbol{i} + 6\boldsymbol{j} - 3\boldsymbol{k})$.

C) Angles (Utilise the dot product): $\boldsymbol{a}.\boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}| \cos\theta$

- Add a modulus sign to the equations below when acute angle is required.

Case 1: Angle between two lines

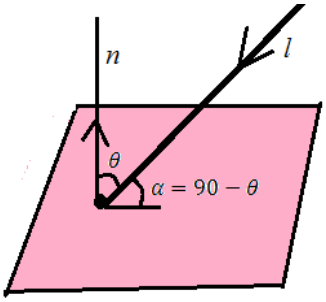


To find the angle between lines AB and BC,

$$\overrightarrow{AB}.\overrightarrow{BC} = |AB||BC| \cos \measuredangle ABC$$

$$\boldsymbol{\cos \measuredangle ABC = \frac{\overrightarrow{AB}.\overrightarrow{BC}}{|AB||BC|}}$$

Case 2: Angle between a line and a plane



Angle between $\pi.\ n = d$ and $r = a + sl, \alpha$:

$$n.l = |n||l| \cos\theta$$

$$\cos\theta = \frac{n.l}{|n||l|}$$

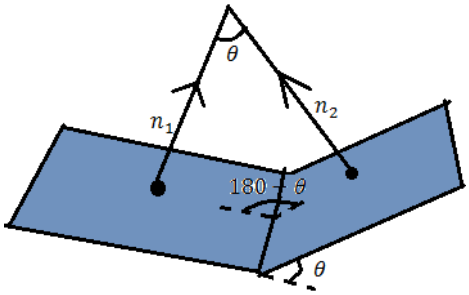
Then $\alpha = 90 - \theta$.

OR:

$$\sin\alpha = \sin(90 - \theta) = \cos\theta$$

$$\boldsymbol{\sin\alpha = \frac{n.l}{|n||l|}}$$

Case 3: Angle between two planes



Angle between $\pi.\ n_1 = d_1$ and $\pi.\ n_2 = d_2$:

$$n_1.n_2 = |n_1||n_2| \cos\theta$$

$$\boldsymbol{\cos\theta = \frac{n_1.n_2}{|n_1||n_2|}}$$

- An easy way to remember this: sine for line and plane and cosine for other two cases.

D) Distances (Master all 6 different situations)

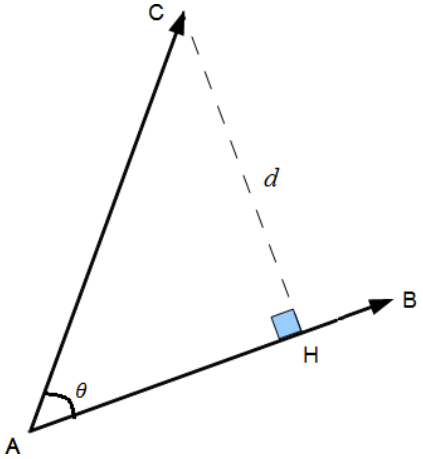
1. Distance from a point to another point

Applying Pythagoras' Theorem, the distance between points $(x_1,\ y_1,\ z_1)$ and $(x_2,\ y_2,\ z_2)$ is

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$$

2. Distance from a point to a line

Task: Find *d*, the distance from point C to line AB.



<u>Method 1: P3 Method</u>

Step 1: Find point H by solving the equation

$$\overrightarrow{CH} \cdot \overrightarrow{AB} = 0$$

$$\boldsymbol{\left(\overrightarrow{OA} + s\overrightarrow{AB} - \overrightarrow{OC}\right) \cdot \overrightarrow{AB} = 0}$$

to find the value of *s*.

Step 2: Substitute *s* into line AB to find $\overrightarrow{OH}$.

Step 3: Find the distance, $d = \left|\overrightarrow{CH}\right|$.

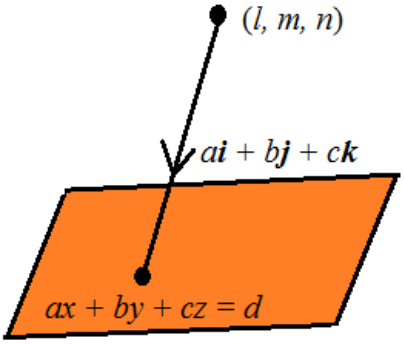
<u>Method 2: FM Method (Cross-product)</u>

We know that

$$\left|\overrightarrow{AB} \times \overrightarrow{AC}\right| = \left|\overrightarrow{AB}\right|\left|\overrightarrow{AC}\right| \sin\theta$$

$$d = \left|\overrightarrow{AC}\right| \sin\theta = \boldsymbol{\frac{\left|\overrightarrow{AB} \times \overrightarrow{AC}\right|}{\left|\overrightarrow{AB}\right|}}.$$

3. Distance from a point to a plane



Form the straight line $\boldsymbol{r} = \begin{pmatrix} l \\ m \\ n \end{pmatrix} + s\begin{pmatrix} a \\ b \\ c \end{pmatrix}$.

When the line intersects the plane,

$$a(l + sa) + b(m + sb) + c(n + sc) = d$$

$$s = \frac{d - al - bm - cn}{a^2 + b^2 + c^2}$$

$$D^2 = \left[\begin{pmatrix} l \\ m \\ n \end{pmatrix} + s\begin{pmatrix} a \\ b \\ c \end{pmatrix} - \begin{pmatrix} l \\ m \\ n \end{pmatrix}\right]^2$$

$$= s^2(a^2 + b^2 + c^2)$$

$$D = |s|\sqrt{a^2 + b^2 + c^2} = \frac{|d - al - bm - cn|}{\sqrt{a^2 + b^2 + c^2}}$$

or more commonly written as

$$\boldsymbol{D = \frac{|al + bm + cn - d|}{\sqrt{a^2 + b^2 + c^2}}}$$

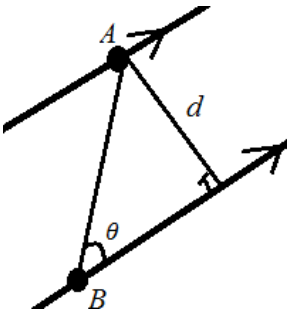
Bonus result: Using the formula above, the distance of the plane $ax + by + cz = d$ from the origin ($l = m = n = 0$) is

$$\frac{|\boldsymbol{d}|}{\sqrt{\boldsymbol{a^2 + b^2 + c^2}}}$$

4. Distance from a line to another line

If the two lines intersect, the distance between them is zero.

If the two lines are parallel,

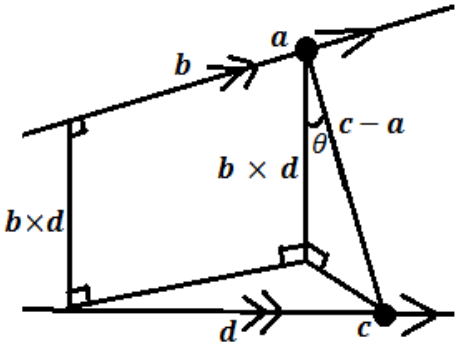


Shortest distance between lines $\overrightarrow{OA} + t\boldsymbol{b}$ and $\overrightarrow{OB} + s\boldsymbol{b}$:

$$\left|\overrightarrow{AB} \times \boldsymbol{b}\right| = \left|\overrightarrow{AB}\right||\boldsymbol{b}| \sin\theta$$

$$d = \left|\overrightarrow{AB}\right| \sin\theta = \frac{\left|\overrightarrow{AB} \times \boldsymbol{b}\right|}{|\boldsymbol{b}|}.$$

If the two lines are skew (New Concept for FM),



Shortest distance between the skew lines $\boldsymbol{a} + t\boldsymbol{b}$ and $\boldsymbol{c} + s\boldsymbol{d}$: (Learn this proof)

$$(\boldsymbol{c} - \boldsymbol{a}) \cdot (\boldsymbol{b} \times \boldsymbol{d}) = |\boldsymbol{c} - \boldsymbol{a}||\boldsymbol{b} \times \boldsymbol{d}| \cos\theta$$

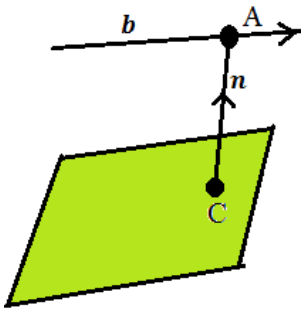
$$\boldsymbol{D = |c - a| \cos\theta = \frac{(c - a) \cdot (b \times d)}{|b \times d|}}$$

- An equation of the common perpendicular is $\boldsymbol{r = a + s(b \times d)}$.

5. Distance from a line to a plane

If the line intersects the plane, the distance between them is 0.

If the line and the plane are parallel:



Distance between the line $\boldsymbol{a} + t\boldsymbol{b}$ and plane π. $n = d$:

Step 1: Form the straight line l: $\boldsymbol{r} = a + t\boldsymbol{n}$

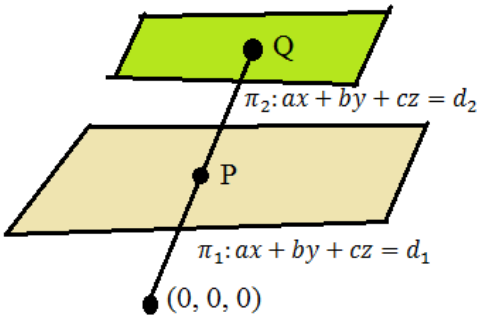
Step 2: Find point C by solving l and π simultaneously then substituting into l.

Step 3: Find the distance, $d = |\overrightarrow{AC}|$.

6. Distance from a plane to another plane

If the planes are not parallel, the distance between them is 0.

If the planes are parallel, they have equations:



Form the line equation $r = (sa, sb, sc)$.

At P,

$$sa^2 + sb^2 + sc^2 = d_1 \rightarrow s = \frac{d_1}{a^2 + b^2 + c^2}$$

At Q,

$$sa^2 + sb^2 + sc^2 = d_2 \rightarrow s = \frac{d_2}{a^2 + b^2 + c^2}$$

$$|PQ|^2 = \left(\frac{d_1 a - d_2 a}{a^2 + b^2 + c^2}\right)^2 + \left(\frac{d_1 b - d_2 b}{a^2 + b^2 + c^2}\right)^2 + \left(\frac{d_1 c - d_2 c}{a^2 + b^2 + c^2}\right)^2 = \frac{(d_1 - d_2)^2}{a^2 + b^2 + c^2}$$

$$\boldsymbol{d = |PQ| = \frac{|d_1 - d_2|}{\sqrt{a^2 + b^2 + c^2}}}$$

8.3 Problem Solving

Example 1: S2003 P1 Q10

Find the acute angle between the planes with equations

$$x - 2y + z - 9 = 0 \quad \text{and} \quad x + y - z + 2 = 0. \qquad [3]$$

The planes meet in the line l, and A is the point on l whose position vector is $p\mathbf{i} + q\mathbf{j} + \mathbf{k}$.

(i) Find p and q. [2]

(ii) Find a vector equation for l. [3]

The non-coincident planes Π_1 and Π_2 are both perpendicular to l. The perpendicular distance from A to Π_1 is $\sqrt{14}$ and the perpendicular distance from A to Π_2 is also $\sqrt{14}$. Find equations for Π_1 and Π_2 in the form $ax + by + cz = d$. [5]

$$n_1 = \boldsymbol{i} - 2\boldsymbol{j} + \boldsymbol{k},\ \ n_2 = \boldsymbol{i} + \boldsymbol{j} - \boldsymbol{k}$$
$$|n_1| = \sqrt{1^2 + 2^2 + 1^2} = \sqrt{6},\quad |n_2| = \sqrt{1^2 + 1^2 + 1^2} = \sqrt{3}$$
$$\cos\theta = \frac{|1 - 2 - 1|}{\sqrt{3}.\sqrt{6}} = \frac{2}{\sqrt{18}} \to \theta = 61.9^o \text{ (to 1 d.p).}$$

(i) $p\boldsymbol{i} + q\boldsymbol{j} + \boldsymbol{k}$ satisfies both equations, thus $p - 2q - 8 = 0 - (1),\ p + q + 1 = 0 - (2)$

$(1) + 2(2)$ gives $3p - 6 = 0 \ \to p = 2, q = -3.$

(ii) $n_1 \times n_2 = (\boldsymbol{i} - 2\boldsymbol{j} + \boldsymbol{k}) \times (\boldsymbol{i} + \boldsymbol{j} - \boldsymbol{k})$

$$= \begin{vmatrix} -2 & 1 \\ 1 & -1 \end{vmatrix}\boldsymbol{i} - \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}\boldsymbol{j} + \begin{vmatrix} 1 & -2 \\ 1 & 1 \end{vmatrix}\boldsymbol{k} = \boldsymbol{i} + 2\boldsymbol{j} + 3\boldsymbol{k}$$

$2\boldsymbol{i} - 3\boldsymbol{j} + \boldsymbol{k}$ is on both planes, thus an equation is $r = 2\boldsymbol{i} - 3\boldsymbol{j} + \boldsymbol{k} + t(\boldsymbol{i} + 2\boldsymbol{j} + 3\boldsymbol{k})$.

π_1 and π_2 are perpendicular to l, thus normal vector of both planes is $\boldsymbol{i} + 2\boldsymbol{j} + 3\boldsymbol{k}$.

Let the equation of the plane be $x + 2y + 3z = d$. Distance from A to π is $\sqrt{14}$:

$$\frac{|2(1) + 2(-3) + 3(1) - d|}{\sqrt{1^2 + 2^2 + 3^2}} = \sqrt{14}$$
$$|-1 - d| = 14$$
$$-1 - d = \pm 14, d = -15, d = 13$$

Thus the planes are $x + 2y + 3z = 13$ and $x + 2y + 3z = -15$.

Example 2: W2005 P1 Q9

The planes Π_1 and Π_2 have vector equations

$$\mathbf{r} = \lambda_1(\mathbf{i} + \mathbf{j} - \mathbf{k}) + \mu_1(2\mathbf{i} - \mathbf{j} + \mathbf{k}) \qquad \text{and} \qquad \mathbf{r} = \lambda_2(\mathbf{i} + 2\mathbf{j} + \mathbf{k}) + \mu_2(3\mathbf{i} + \mathbf{j} - \mathbf{k})$$

respectively. The line l passes through the point with position vector $4\mathbf{i} + 5\mathbf{j} + 6\mathbf{k}$ and is parallel to both Π_1 and Π_2. Find a vector equation for l. [6]

Find also the shortest distance between l and the line of intersection of Π_1 and Π_2. [4]

$$n_1 = (\boldsymbol{i} + \boldsymbol{j} - \boldsymbol{k}) \times (2\boldsymbol{i} - \boldsymbol{j} + \boldsymbol{k})$$
$$= \begin{vmatrix} 1 & -1 \\ -1 & 1 \end{vmatrix}\boldsymbol{i} - \begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix}\boldsymbol{j} + \begin{vmatrix} 1 & 1 \\ 2 & -1 \end{vmatrix}\boldsymbol{k}$$
$$= -3\boldsymbol{j} - 3\boldsymbol{k}$$

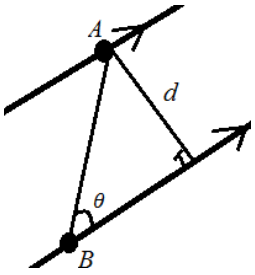
$$n_2 = (\boldsymbol{i} + 2\boldsymbol{j} + \boldsymbol{k}) \times (3\boldsymbol{i} + \boldsymbol{j} - \boldsymbol{k})$$
$$= \begin{vmatrix} 2 & 1 \\ 1 & -1 \end{vmatrix}\boldsymbol{i} - \begin{vmatrix} 1 & 1 \\ 3 & -1 \end{vmatrix}\boldsymbol{j} + \begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix}\boldsymbol{k}$$
$$= -3\boldsymbol{i} + 4\boldsymbol{j} - 5\boldsymbol{k}$$

$$n_1 \times n_2 = \begin{vmatrix} -3 & -3 \\ 4 & -5 \end{vmatrix}\boldsymbol{i} - \begin{vmatrix} 0 & -3 \\ -3 & -5 \end{vmatrix}\boldsymbol{j} + \begin{vmatrix} 0 & -3 \\ -3 & 4 \end{vmatrix}\boldsymbol{k} = 27\boldsymbol{i} + 9\boldsymbol{j} - 9\boldsymbol{k}$$

The direction vector of l is $3\boldsymbol{i} + \boldsymbol{j} - \boldsymbol{k}$. An equation of l is $r = 4\boldsymbol{i} + 5\boldsymbol{j} + 6\boldsymbol{k} + s(3\boldsymbol{i} + \boldsymbol{j} - \boldsymbol{k})$.

The line of intersection has equation $r = t(3\boldsymbol{i} + \boldsymbol{j} - \boldsymbol{k})$.

$$\overrightarrow{AB} = 4\boldsymbol{i} + 5\boldsymbol{j} + 6\boldsymbol{k}$$
$$\overrightarrow{AB} \times n = \begin{vmatrix} 5 & 6 \\ 1 & -1 \end{vmatrix}\boldsymbol{i} - \begin{vmatrix} 4 & 6 \\ 3 & -1 \end{vmatrix}\boldsymbol{j} + \begin{vmatrix} 4 & 5 \\ 3 & 1 \end{vmatrix}\boldsymbol{k} = -11\boldsymbol{i} + 22\boldsymbol{j} - 11\boldsymbol{k}$$
$$|\overrightarrow{AB} \times n| = |\overrightarrow{AB}||n| \sin\theta$$
$$d = |\overrightarrow{AB}| \sin\theta = \frac{\sqrt{11^2 + 22^2 + 11^2}}{\sqrt{3^2 + 1^2 + 1^2}} = \sqrt{66}$$



Example 3: S2006 P1 Q10

The equation of the plane Π is

$$2x + 3y + 4z = 48.$$

Obtain a vector equation of Π in the form

$$\mathbf{r} = \mathbf{a} + \lambda\mathbf{b} + \mu\mathbf{c},$$

where **a**, **b** and **c** are of the form $p\mathbf{i}$, $q\mathbf{i} + r\mathbf{j}$ and $s\mathbf{i} + t\mathbf{k}$ respectively, and where p, q, r, s, t are integers. [6]

The line l has vector equation $\mathbf{r} = 29\mathbf{i} - 2\mathbf{j} - \mathbf{k} + \theta(5\mathbf{i} - 6\mathbf{j} + 2\mathbf{k})$. Show that l lies in Π. [3]

Find, in the form $ax + by + cz = d$, the equation of the plane which contains l and is perpendicular to Π. [4]

$$\begin{aligned} 2\boldsymbol{i} + 3\boldsymbol{j} + 4\boldsymbol{k} &= (q\boldsymbol{i} + r\boldsymbol{j}) \times (s\boldsymbol{i} + t\boldsymbol{k}) \\ &= \begin{vmatrix} r & 0 \\ 0 & t \end{vmatrix}\boldsymbol{i} - \begin{vmatrix} q & 0 \\ s & t \end{vmatrix}\boldsymbol{j} + \begin{vmatrix} q & r \\ s & 0 \end{vmatrix}\boldsymbol{k} \\ &= (rt)\boldsymbol{i} - (qt)\boldsymbol{j} - (rs)\boldsymbol{k} \end{aligned}$$

$$\begin{aligned} (p\boldsymbol{i}) \cdot (2\boldsymbol{i} + 3\boldsymbol{j} + 4\boldsymbol{k}) &= 48 \\ 2p &= 48 \\ p &= 24 \end{aligned}$$

As q, r, s, t are integers, possible values are $t = 1, r = -2, q = 3, s = 2$.

A vector equation for π is $r = 24\boldsymbol{i} + \lambda(3\boldsymbol{i} - 2\boldsymbol{j}) + \mu(2\boldsymbol{i} - \boldsymbol{k})$.

Points in l have coordinates $(29 + 5\theta, -2 - 6\theta, -1 + 2\theta)$

When l meets π: LHS: $2(29 + 5\theta) + 3(-2 - 6\theta) + 4(-1 + 2\theta) = 48 = RHS$

Since every point in l is in π, l lies in π.

2 vectors in the plane are $2\boldsymbol{i} + 3\boldsymbol{j} + 4\boldsymbol{k}$ and $5\boldsymbol{i} - 6\boldsymbol{j} + 2\boldsymbol{k}$.

$$(2\boldsymbol{i} + 3\boldsymbol{j} + 4\boldsymbol{k}) \times (5\boldsymbol{i} - 6\boldsymbol{j} + 2\boldsymbol{k}) = \begin{vmatrix} 3 & 4 \\ -6 & 2 \end{vmatrix}\boldsymbol{i} - \begin{vmatrix} 2 & 4 \\ 5 & 2 \end{vmatrix}\boldsymbol{j} + \begin{vmatrix} 2 & 3 \\ 5 & -6 \end{vmatrix}\boldsymbol{k} = 30\boldsymbol{i} + 16\boldsymbol{j} - 27\boldsymbol{k}$$

A point in the plane is $29\boldsymbol{i} - 2\boldsymbol{j} - \boldsymbol{k}$: $d = 29(30) - 16(2) + 27(1) = 865$.

Thus the equation of the plane is $30x + 16y - 27z = 865$.

Example 4: S2008 P1 Q12 EITHER

The position vectors of the points A, B, C, D are

$$7\mathbf{i} + 4\mathbf{j} - \mathbf{k}, \qquad 3\mathbf{i} + 5\mathbf{j} - 2\mathbf{k}, \qquad 2\mathbf{i} + 6\mathbf{j} + 3\mathbf{k}, \qquad 2\mathbf{i} + 7\mathbf{j} + \lambda\mathbf{k}$$

respectively. It is given that the shortest distance between the line AB and the line CD is 3.

(i) Show that $\lambda^2 - 5\lambda + 4 = 0$. [7]

(ii) Find the acute angle between the planes through A, B, D corresponding to the values of λ satisfying the equation in part **(i)**. [7]

(i)

$$\overrightarrow{BA} = (7\boldsymbol{i} + 4\boldsymbol{j} - \boldsymbol{k}) - (3\boldsymbol{i} + 5\boldsymbol{j} - 2\boldsymbol{k}) = 4\boldsymbol{i} - \boldsymbol{j} + \boldsymbol{k}$$

$$\overrightarrow{DC} = (2\boldsymbol{i} + 7\boldsymbol{j} + \lambda\boldsymbol{k}) - (2\boldsymbol{i} + 6\boldsymbol{j} + 3\boldsymbol{k}) = \boldsymbol{j} + (\lambda - 3)\boldsymbol{k}$$

$$\begin{aligned} n = \overrightarrow{BA} \times \overrightarrow{DC} &= \begin{vmatrix} -1 & 1 \\ 1 & \lambda - 3 \end{vmatrix}\boldsymbol{i} - \begin{vmatrix} 4 & 1 \\ 0 & \lambda - 3 \end{vmatrix}\boldsymbol{j} + \begin{vmatrix} 4 & -1 \\ 0 & 1 \end{vmatrix}\boldsymbol{k} \\ &= (2 - \lambda)\boldsymbol{i} - 4(\lambda - 3)\boldsymbol{j} + 4\boldsymbol{k} \end{aligned}$$

$$\overrightarrow{CA} = (7\boldsymbol{i} + 4\boldsymbol{j} - \boldsymbol{k}) - (2\boldsymbol{i} + 6\boldsymbol{j} + 3\boldsymbol{k}) = 5\boldsymbol{i} - 2\boldsymbol{j} - 4\boldsymbol{k}$$

$$\overrightarrow{CA} \cdot n = 5(2 - \lambda) + 8(\lambda - 3) - 16 = 3\lambda - 30$$

$$|n| = \sqrt{(2 - \lambda)^2 + 16(\lambda - 3)^2 + 4^2} = \sqrt{17\lambda^2 - 100\lambda + 64}$$

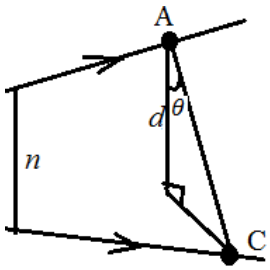
$$\left|\overrightarrow{CA}\right||n|\cos\theta = \overrightarrow{CA}\cdot n;\ d = \left|\overrightarrow{CA}\right|\cos\theta = 3:$$

$$3\sqrt{17\lambda^2+100\lambda+164} = 3\lambda - 30$$

$$17\lambda^2 - 100\lambda + 164 = (\lambda-10)^2$$

$$16\lambda^2 - 80\lambda + 64 = 0$$

$$\lambda^2 - 5\lambda + 4 = 0$$



(ii) $\lambda^2 - 5\lambda + 4 = 0;\ \lambda = 1, \lambda = 4.$

$$\overrightarrow{DA} = (7\boldsymbol{i}+4\boldsymbol{j}-\boldsymbol{k}) - (2\boldsymbol{i}+7\boldsymbol{j}+\lambda\boldsymbol{k})$$
$$= 5\boldsymbol{i} - 3\boldsymbol{j} - (1+\lambda)\boldsymbol{k}$$
$$\overrightarrow{BA}\times\overrightarrow{DA}$$
$$= \begin{vmatrix}-1 & 1\\ -3 & -(1+\lambda)\end{vmatrix}\boldsymbol{i} - \begin{vmatrix}4 & 1\\ 5 & -(1+\lambda)\end{vmatrix}\boldsymbol{j} + \begin{vmatrix}4 & -1\\ 5 & -3\end{vmatrix}\boldsymbol{k}$$
$$= (4+\lambda)\boldsymbol{i} + (9+4\lambda)\boldsymbol{j} - 7\boldsymbol{k}$$

$$\lambda = 1: m_1 = 5\boldsymbol{i} + 13\boldsymbol{j} - 7\boldsymbol{k}$$
$$\lambda = 4: m_4 = 8\boldsymbol{i} + 25\boldsymbol{j} - 7\boldsymbol{k}$$
$$|m_1| = \sqrt{5^2+13^2+7^2} = \sqrt{243}$$
$$|m_4| = \sqrt{8^2+25^2+7^2} = \sqrt{738}$$
$$m_1\cdot m_4 = 40 + 325 + 49 = 414$$
$$\cos\theta = \frac{414}{\sqrt{243}\sqrt{738}} \rightarrow \theta = 12.1^o$$

Example 5: S2009 P1 Q11

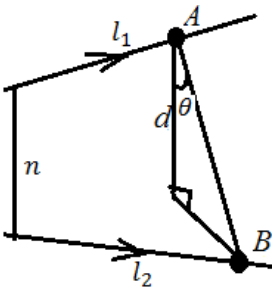
The line l_1 is parallel to the vector $4\mathbf{j} - \mathbf{k}$ and passes through the point A whose position vector is $2\mathbf{i} + \mathbf{j} + 4\mathbf{k}$. The variable line l_2 is parallel to the vector $\mathbf{i} - (2\sin t)\mathbf{j}$, where $0 \leqslant t < 2\pi$, and passes through the point B whose position vector is $\mathbf{i} + 2\mathbf{j} + 4\mathbf{k}$. The points P and Q are on l_1 and l_2, respectively, and PQ is perpendicular to both l_1 and l_2.

(i) Find the length of PQ in terms of t. [5]

(ii) Hence find the values of t for which l_1 and l_2 intersect. [2]

(iii) For the case $t = \frac{1}{4}\pi$, find the perpendicular distance from A to the plane BPQ, giving your answer correct to 3 decimal places. [5]

(i)
$$n = \begin{vmatrix}4 & -1\\ -2\sin t & 0\end{vmatrix}\boldsymbol{i} - \begin{vmatrix}0 & -1\\ 1 & 0\end{vmatrix}\boldsymbol{j} + \begin{vmatrix}0 & 4\\ 1 & -2\sin t\end{vmatrix}\boldsymbol{k}$$
$$= -2\sin t\,\boldsymbol{i} - \boldsymbol{j} - 4\boldsymbol{k}$$
$$\overrightarrow{AB} = (\boldsymbol{i}+2\boldsymbol{j}+4\boldsymbol{k}) - (2\boldsymbol{i}+\boldsymbol{j}+4\boldsymbol{k}) = -\boldsymbol{i}+\boldsymbol{j}$$
$$\left|\overrightarrow{AB}\cdot n\right| = \left|\overrightarrow{AB}\right||n|\cos\theta$$
$$PQ = \left|\overrightarrow{AB}\right|\cos\theta = \frac{\left|\overrightarrow{AB}\cdot n\right|}{|n|} = \frac{|2\sin t - 1|}{\sqrt{4sin^2t+17}}$$



(ii) When the two lines intersect, $PQ = 0$.

$$2\sin t - 1 = 0, \sin t = \frac{1}{2}, t = \frac{1}{6}\pi \text{ or } \frac{5}{6}\pi$$

(iii) $t = \frac{1}{4}\pi$: $\overrightarrow{PQ} = n = -\sqrt{2}\boldsymbol{i} - \boldsymbol{j} - 4\boldsymbol{k}$

Both B and Q are on l_2: $\overrightarrow{BQ} = \boldsymbol{i} - \sqrt{2}\boldsymbol{j}$

$$\overrightarrow{BQ}\times\overrightarrow{PQ} = \begin{vmatrix}-\sqrt{2} & 0\\ -1 & -4\end{vmatrix}\boldsymbol{i} - \begin{vmatrix}1 & 0\\ -\sqrt{2} & -4\end{vmatrix}\boldsymbol{j} + \begin{vmatrix}1 & -\sqrt{2}\\ -\sqrt{2} & -1\end{vmatrix}\boldsymbol{k} = 4\sqrt{2}\boldsymbol{i} + 4\boldsymbol{j} - 3\boldsymbol{k}$$
$$d = \left(4\sqrt{2}\boldsymbol{i} + 4\boldsymbol{j} - 3\boldsymbol{k}\right)\cdot(\boldsymbol{i}+2\boldsymbol{j}+4\boldsymbol{k}) = 4\sqrt{2} - 4$$

Plane BPQ has equation $4\sqrt{2}x + 4y - 3z = 4\sqrt{2} - 4$

$$D = \frac{\left|4\sqrt{2}(2)+4(1)-3(4)-4\sqrt{2}+4\right|}{\sqrt{\left(4\sqrt{2}\right)^2+4^2+3^2}} = \frac{4\sqrt{2}-4}{\sqrt{57}} \approx 0.219$$

CHAPTER 9: MATRICES AND LINEAR SPACES

9.1 Pre-requisite Knowledge: Matrix Algebra

- Note: This section covers pre-requisite contents from IGCSE Additional Mathematics or any equivalent qualifications.

A) Matrix Definitions

- A matrix is a rectangular array of arrangements of quantities in rows and columns, and can be used to represent quantities like vectors, transformations and equations.
- A matrix with *m* rows and *n* columns has order $m \times n$.
- Every entry is known as an element, and an element in row *m* and column *n* is denoted as a_{mn}.
- A row matrix is a matrix with only one row, i.e. order $1 \times n$.
- A column matrix is a matrix with only one column, i.e. order $n \times 1$.
- A null matrix is a matrix with '0's as all of its elements.
- A square matrix is a matrix with equal number of rows and columns, i.e. order $m \times m$.
- A diagonal matrix is a matrix (usually square matrix) with elements only in its diagonal and '0's as all other elements
- An identity matrix, **I** is a square, diagonal matrix with '1's in the diagonal and '0's as all other elements.
- Examples:

A 2×4 matrix: $\begin{bmatrix} 1 & -3 & 5 & 0 \\ 0 & 6 & 9 & -2 \end{bmatrix}$ $a_{22} = 6, a_{13} = 5$	A 3×3 (square) matrix: $\begin{bmatrix} 1 & -4 & 6 \\ 2 & -6 & -1 \\ 5 & 0 & 0 \end{bmatrix}$
A row matrix (1×3): $\begin{bmatrix} -5 & 6 & 7 \end{bmatrix}$	A column matrix (4×1): $\begin{bmatrix} 0 \\ 2 \\ -1 \\ 0 \end{bmatrix}$
A null matrix (2×2): $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$	A 2×2 identity matrix: $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
A diagonal matrix (3×3): $\begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{bmatrix}$	A 3×3 identity matrix: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

B) Operations of Matrices

Addition	$\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\\7 & 8 & 9\end{bmatrix}+\begin{bmatrix}6 & 0 & -3\\4 & 2 & 0\\-1 & 5 & -2\end{bmatrix}=\begin{bmatrix}1+6 & 2+0 & 3-3\\4+4 & 5+2 & 6+0\\7-1 & 8+5 & 9-2\end{bmatrix}=\begin{bmatrix}7 & 2 & 0\\8 & 7 & 6\\6 & 13 & 7\end{bmatrix}$
Subtraction	$\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\\7 & 8 & 9\end{bmatrix}-\begin{bmatrix}6 & 0 & -3\\4 & 2 & 0\\-1 & 5 & -2\end{bmatrix}=\begin{bmatrix}1-6 & 2-0 & 3+3\\4-4 & 5-2 & 6-0\\7+1 & 8-5 & 9+2\end{bmatrix}=\begin{bmatrix}-5 & 2 & 6\\0 & 3 & 6\\8 & 3 & 11\end{bmatrix}$
Scalar Multiplication	$a\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\\7 & 8 & 9\end{bmatrix}=\begin{bmatrix}a & 2a & 3a\\4a & 5a & 6a\\7a & 8a & 9a\end{bmatrix}$
Matrix Multiplication	$\begin{bmatrix}1 & 4 & 7\\2 & 5 & 8\\3 & 6 & 9\end{bmatrix}\begin{bmatrix}a & d & g\\b & e & h\\c & f & i\end{bmatrix}$ $=\begin{bmatrix}1(a)+4(b)+7(c) & 1(d)+4(e)+7(f) & 1(g)+4(h)+7(i)\\2(a)+5(b)+8(c) & 2(d)+5(e)+8(f) & 2(g)+5(h)+8(i)\\3(a)+6(b)+9(c) & 3(d)+6(e)+9(f) & 3(g)+6(h)+9(i)\end{bmatrix}$
Transpose of a matrix	$A=\begin{bmatrix}a & d & g\\b & e & h\\c & f & i\end{bmatrix}$, then $A^T=\begin{bmatrix}a & d & g\\b & e & h\\c & f & i\end{bmatrix}$,

- Multiplication of a matrix is done using the 'row by column' addition method.
- Division of a matrix with another matrix could not be performed.
- Transpose of a matrix is obtained by interchanging its rows and columns.
- Addition and subtraction of matrices can only be performed if their orders are equal.
- For multiplication of two matrices **A** and **B**, $\mathbf{AB} \neq \mathbf{BA}$ in general. In this chapter, we will indicate matrix multiplication using the terms 'pre-multiply' and 'post-multiply'.
- Multiplication of a matrix **A** with an identity matrix **I** equals itself if they have the same order, i.e. $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$.
- As multiplication of a matrix is defined as multiplication of row by column, it can only be carried out if the number of column in the first matrix equals the number of rows of the second matrix, as shown below:

If **A** is a $a \times b$ matrix, and:

1. **B** is a $b \times c$ matrix, then **AB** is a $a \times c$ matrix and **BA** does not exist;
2. **B** is a $a \times c$ matrix, then both **AB** and **BA** do not exist;
3. **B** is a $b \times a$ matrix, then **AB** is a $a \times a$ matrix and **BA** is a $b \times b$ matrix.

9.2 Determinants and Inverses

- LO: evaluate the determinant of a square matrix and find the inverse of a non-singular matrix (2×2 and 3×3 matrices only);

A) Determinant of a Matrix

- Determinants are only defined for square matrices. Determinant of matrix **A** is denoted as $\det(\mathbf{A})$.
- If $\det(\mathbf{A}) = 0$, matrix **A** is said to be singular.
- Determinants of 2×2 and 3×3 matrices are defined as of follows:

For a 2×2 matrix $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$:	$\det(\boldsymbol{A}) = ad - bc$
For a 3×3 matrix $\boldsymbol{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$:	$\det(\boldsymbol{A}) = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}$ OR $\det(\boldsymbol{A}) = -d\begin{vmatrix} b & c \\ h & i \end{vmatrix} + e\begin{vmatrix} a & c \\ g & i \end{vmatrix} - f\begin{vmatrix} a & b \\ g & h \end{vmatrix}$ OR $\det(\boldsymbol{A}) = g\begin{vmatrix} b & c \\ e & f \end{vmatrix} - h\begin{vmatrix} a & c \\ d & f \end{vmatrix} + i\begin{vmatrix} a & b \\ d & e \end{vmatrix}$ (all three formulas yield the same result)

- For a 3 by 3 matrix, we can expand along any row, but usually we pick the row with the most zeroes to save time.

B) Inverse Matrices

- The inverse matrix of **A** is denoted as $\mathbf{A}^{-1}$, and are related by $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$.
- If matrices **A** and **B** have the same order, then $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

Proof: Let $\mathbf{C} = \mathbf{AB}$. We have $\mathbf{B} = \mathbf{A}^{-1}\mathbf{AB} = \mathbf{A}^{-1}\mathbf{C}$ and $\mathbf{A} = \mathbf{ABB}^{-1} = \mathbf{CB}^{-1}$. Thus

$$\mathbf{AB} = \mathbf{CB^{-1}A^{-1}C}$$
$$\mathbf{C} = \mathbf{CB^{-1}A^{-1}C}$$
$$\mathbf{C^{-1}CC^{-1}} = \mathbf{C^{-1}CB^{-1}A^{-1}CC^{-1}}$$
$$\mathbf{IC^{-1}} = \mathbf{IB^{-1}A^{-1}I}$$
$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$

- The inverse of a *n* by *n* matrix is defined as of follows, this also implies that the inverse of a matrix does not exist (singular) if its determinant is zero.

For a general $n \times n$ matrix **A**:	$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})}\, adj\,(\mathbf{A})$
Applying this rule to a 2×2 matrix $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$:	$\mathbf{A}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

- In the following example, we will use this rule to find the inverse of a 3 by 3 matrix.

Example: Given $\mathbf{A} = \begin{bmatrix} 1 & 4 & -1 \\ -2 & 1 & 2 \\ 3 & 5 & -1 \end{bmatrix}$, find $\boldsymbol{A}^{-1}$.

Step 1: Determinant	$\det(\boldsymbol{A}) = \begin{vmatrix} 1 & 2 \\ 5 & -1 \end{vmatrix} - 4\begin{vmatrix} -2 & 2 \\ 3 & -1 \end{vmatrix} - \begin{vmatrix} -2 & 1 \\ 3 & 5 \end{vmatrix} = -11 + 16 + 13 = 18$
Step 2: Minors	Matrix of minors of a 3×3 matrix $\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$ is as follows: $\begin{bmatrix} \begin{vmatrix} e & f \\ h & i \end{vmatrix} & \begin{vmatrix} d & f \\ g & i \end{vmatrix} & \begin{vmatrix} d & e \\ g & h \end{vmatrix} \\ \begin{vmatrix} b & c \\ h & i \end{vmatrix} & \begin{vmatrix} a & c \\ g & i \end{vmatrix} & \begin{vmatrix} a & b \\ g & h \end{vmatrix} \\ \begin{vmatrix} b & c \\ e & f \end{vmatrix} & \begin{vmatrix} a & c \\ d & f \end{vmatrix} & \begin{vmatrix} a & b \\ d & e \end{vmatrix} \end{bmatrix}$ Thus $\begin{bmatrix} \begin{vmatrix} 1 & 2 \\ 5 & -1 \end{vmatrix} & \begin{vmatrix} -2 & 2 \\ 3 & -1 \end{vmatrix} & \begin{vmatrix} -2 & 1 \\ 3 & 5 \end{vmatrix} \\ \begin{vmatrix} 4 & -1 \\ 5 & -1 \end{vmatrix} & \begin{vmatrix} 1 & -1 \\ 3 & -1 \end{vmatrix} & \begin{vmatrix} 1 & 4 \\ 3 & 5 \end{vmatrix} \\ \begin{vmatrix} 4 & -1 \\ 1 & 2 \end{vmatrix} & \begin{vmatrix} 1 & -1 \\ -2 & 2 \end{vmatrix} & \begin{vmatrix} 1 & 4 \\ -2 & 1 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} -11 & -4 & -13 \\ 1 & 2 & -7 \\ 9 & 0 & 9 \end{bmatrix}$
Step 3: Cofactors	Change the signs of the minor matrix as follows: For 2×2 matrix: $\begin{bmatrix} + & - \\ - & + \end{bmatrix}$, For 3×3 matrix: $\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$ '+' sign indicates that the element is unchanged, while '–' sign indicates that the element is multiplied by –1 . Thus $\begin{bmatrix} -11 & -4 & -13 \\ 1 & 2 & -7 \\ 9 & 0 & 9 \end{bmatrix} \rightarrow \begin{bmatrix} -11 & 4 & -13 \\ -1 & 2 & 7 \\ 9 & 0 & 9 \end{bmatrix}$
Step 4: Adjoint	Adjoint matrix is obtained by transposing the cofactor matrix. $adj\ (\mathbf{A}) = \begin{bmatrix} -11 & 4 & -13 \\ -1 & 2 & 7 \\ 9 & 0 & 9 \end{bmatrix}^T = \begin{bmatrix} -11 & -1 & 9 \\ 4 & 2 & 0 \\ -13 & 7 & 9 \end{bmatrix}$
Step 5: Inverse	$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})}\ adj\ (\mathbf{A}) = \frac{1}{18}\begin{bmatrix} -11 & -1 & 9 \\ 4 & 2 & 0 \\ -13 & 7 & 9 \end{bmatrix}$ Check your answer: multiply it with the original matrix and make sure that you get **I**.

- Alternative method: make three column vectors from the 3 columns (i.e. $\boldsymbol{a} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} 4 \\ 1 \\ 5 \end{bmatrix}$, $\boldsymbol{c} = \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}$ from example above). The cofactor matrix in Step 3 is then $[(\boldsymbol{b} \times \boldsymbol{c}) \quad (\boldsymbol{c} \times \boldsymbol{a}) \quad (\boldsymbol{a} \times \boldsymbol{b})]$, combining Steps 2 and 3 above together.

9.3 System of Linear Equations

- LO: use methods associated with matrices in the context of the solution of a set of linear equations;

A) Matrices and Equations

- A linear equation is in the form $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$. All unknowns are only raised to power one. It is homogenous if $b = 0$ and inhomogeneous otherwise.
- A system of linear equations can be represented by a matrix as seen below:

$2x - 3y = 4$ $5x + 2y = 6$	$\begin{bmatrix} 2 & -3 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$
$2x_1 - 6x_2 + x_4 = 0$ $x_1 + x_2 + 3x_3 - 2x_4 = 5$	$\begin{bmatrix} 2 & -6 & 0 & 1 \\ 1 & 1 & 3 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}$

- The general form of a system of linear equations is $\mathbf{Ax} = \mathbf{b}$. If $\mathbf{A}^{-1}$ exists, we can pre-multiply the equation with $\mathbf{A}^{-1}$:

$$\mathbf{Ax} = \mathbf{b}$$
$$\mathbf{A}^{-1}\,\mathbf{Ax} = \mathbf{A}^{-1}\,\mathbf{b}$$
$$\mathbf{Ix} = \mathbf{A}^{-1}\,\mathbf{b} \Rightarrow \mathbf{x} = \mathbf{A}^{-1}\,\mathbf{b}$$

- Example: Solve the system of equations below:

$$x + 4y - z = 1 \quad -2x + y + 2z = 2 \quad 3x + 5y - z = 5$$

Writing the system in matrix form, let $\mathbf{A} = \begin{bmatrix} 1 & 4 & -1 \\ -2 & 1 & 2 \\ 3 & 5 & -1 \end{bmatrix}$: $\mathbf{A}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$

Pre-multiplying both sides by $\mathbf{A}^{-1}$ (which is found in the previous section):

$$\mathbf{A}^{-1}\mathbf{A}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{A}^{-1}\begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$$
$$\mathbf{I}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{18}\begin{bmatrix} -11 & -1 & 9 \\ 4 & 2 & 0 \\ -13 & 7 & 9 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$$
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{18}\begin{bmatrix} -11 - 2 + 45 \\ 4 + 4 \\ -13 + 14 + 45 \end{bmatrix} = \frac{1}{18}\begin{bmatrix} 32 \\ 8 \\ 46 \end{bmatrix}$$

Giving us $x = \frac{16}{9}$, $y = \frac{4}{9}$, $z = \frac{23}{9}$.

B) Elementary Row Operations (ERO) / Gauss-Jordan Elimination

- The inverse matrix method cannot solve systems that have a matrix with determinant zero. To solve these types of equations, we can write the system into an augmented matrix and reduce it using elementary row operations. For example:

$5x + 6y - 3z = 8$ $y - 2z = 1$ $ax + 5y - az = 0$	$\left(\begin{array}{ccc\|c} 5 & 6 & -3 & 8 \\ 0 & 1 & -2 & 1 \\ a & 5 & -a & 0 \end{array}\right)$
$2x_1 - 6x_2 + x_4 = 0$ $x_1 + x_2 + 3x_3 - 2x_4 = 5$	$\left(\begin{array}{cccc\|c} 2 & -6 & 0 & 1 & 0 \\ 1 & 1 & 3 & -2 & 5 \end{array}\right)$

- Elementary row operations are a set of three rules that we apply to each element to a row of the matrix. The three rules are:

1. Multiply a row through by a nonzero constant.
2. Interchange any two rows.
3. Multiply a row with a constant then add the result to another row.

- These operations can reduce an augmented matrix into its reduced row-echelon form, which satisfies the following characteristics:

1. All zero rows are at the bottom of the matrix
2. The leading entry of each nonzero row after the first occurs to the right of the leading entry of the previous row.
3. The leading entry in any nonzero row is 1.
4. All entries in the column above and below a leading 1 are zero.

- The following example shows the two methods to solve a system of equations.

Normal Method	**Elementary Row Operations**
$x+y+2z=9$ $2x+4y-3z=1$ $3x+6y-5z=0$	Augmented matrix: $\left(\begin{array}{ccc\|c} 1 & 1 & 2 & 9 \\ 2 & 4 & -3 & 1 \\ 3 & 6 & -5 & 0 \end{array}\right)$
Add -2 times the 1st equation to the 2nd: Add -3 times the 1st equation to the 3rd:	Add -2 times the first row to the second: Add -3 times the first row to the third :
$x+y+2z=9$ $2y-7z=-17$ $3y-11z=-27$	$\xrightarrow{\substack{-2R_1+R_2\to R_2 \\ -3R_1+R_3\to R_3}}$ $\left(\begin{array}{ccc\|c} 1 & 1 & 2 & 9 \\ 0 & 2 & -7 & -17 \\ 0 & 3 & -11 & -27 \end{array}\right)$
Multiply the second equation by ½: Add -3 times the 2nd equation to the 3rd : Multiply the third equation by -2:	Multiply the second row by ½: Add -3 times the second row to the third: Multiply the third row by -2 to obtain:
$x+y+2z=9$ $y-\frac{7}{2}z=-\frac{17}{2}$ $z=3$	$\xrightarrow{\substack{\frac{1}{2}R_2\to R_2 \\ -3R_2+R_3\to R_3 \\ 2R_3\to R_3}}$ $\left(\begin{array}{ccc\|c} 1 & 1 & 2 & 9 \\ 0 & 1 & -\frac{7}{2} & -\frac{17}{2} \\ 0 & 0 & 1 & 3 \end{array}\right)$
Add -1 times the 2nd equation to the 1st: Add $-\frac{11}{2}$ times the 3rd equation to the 1st: Add $\frac{7}{2}$ times the 3rd equation to the 2nd:	Add -1 times the second row to the first: Add $-\frac{11}{2}$ times the third row to the first: Add $\frac{7}{2}$ times the third row to the second:
$x=1$ $y=2$ $z=3$	$\xrightarrow{\substack{-R_2+R_1\to R_1 \\ -\frac{11}{2}R_3+R_1\to R_1 \\ \frac{7}{2}R_3+R_2\to R_2}}$ $\left(\begin{array}{ccc\|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right)$

- From the above example, we can see that performing elementary row operations on a matrix is equivalent to solving the system.
- Note: Learn how elementary row operations are denoted on the example above. We will use it a lot in the future parts of this topic as the systems we have to solve get much complex than what we solved above.
- Another application of ERO: to find inverse matrix

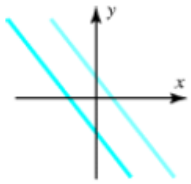
To find the inverse of $\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$, write it in row-echelon form

$$\left(\begin{array}{ccc\|ccc} a & b & c & 1 & 0 & 0 \\ d & e & f & 0 & 1 & 0 \\ g & h & i & 0 & 0 & 1 \end{array}\right)$$

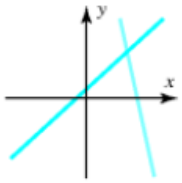
Then perform elementary row operations until the matrix on the left is reduced to the identity matrix. The matrix that is produced at the right is the answer.

C) Consistency of Equations

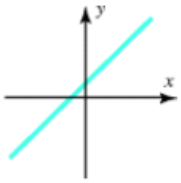
- A system of equations is called consistent if there is at least one solution. In this case the system may have one solution or infinite solutions.
- If there are no solutions to the system it is known to be inconsistent.
- Generally, only system of *n* linear equation in *n* variables will have a unique solution.
- Also when a system of linear equations is written in the form $\mathbf{Ax} = \mathbf{b}$, $\det(\mathbf{A})$ cannot equal 0 for the system to have a unique solution.
- A system of linear equations in two variables (2 equations in the form $a_1x_1 + a_2x_2 = b$) are a pair of straight lines. Only 3 possible cases exist:



No solution



One solution



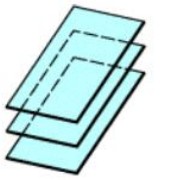
Infinitely many solutions (coincident lines)

$x + y = 3 - (1)$
$3x + 3y = 6 - (2)$
$(1) \times 3: 9 = 6$
No solutions.

$x - y = 1 - (1)$
$x + y = 5 - (2)$
$(1) + (2): 2x = 6$
$x = 3, y = 5 - 3 = 2$

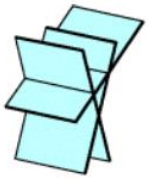
$x - y = 1 - (1)$
$4x - 4y = 4 - (2)$
$(1) \times 4 = (2)$
Let $x = t, y = t - 1$.
The solution is a line.

- Remember from the last chapter that a system of linear equations in three variables represent the equation of multiple planes. Many cases can occur as shown below:

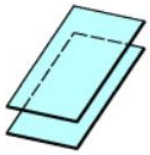


No solutions (three parallel planes; no common intersection)

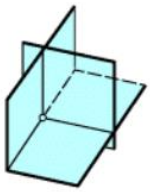
No solutions (two parallel planes; no common intersection)



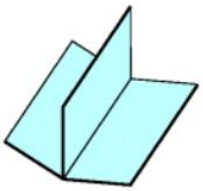
No solutions (no common intersection)



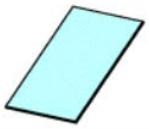
No solutions (two coincident planes parallel to the third; no common intersection)



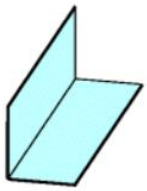
One solution (intersection is a point)



Infinitely many solutions (intersection is a line)



Infinitely many solutions (planes are all coincident; intersection is a plane)



Infinitely many solutions (two coincident planes; intersection is a line)

- For systems with unknowns as coefficients, the number of solutions to the system will vary, as shown below:

Example 1: Solve the system of equations below:

$$ax + y + z = 1 \qquad x + ay + z = a \qquad x + y + az = a^2$$

Start by interchanging rows (1) and (2) by *a* and writing the system in row-echelon form:

$$\left(\begin{array}{ccc|c}1 & a & 1 & a\\ a & 1 & 1 & 1\\ 1 & 1 & a & a^2\end{array}\right) \xrightarrow{\substack{-aR_1+R_2\to R_2\\ -R_1+R_3\to R_3}} \left(\begin{array}{ccc|c}1 & a & 1 & a\\ 0 & 1-a^2 & 1-a & 1-a^2\\ 0 & 1-a & a-1 & a^2-a\end{array}\right)$$

$$\xrightarrow{\substack{R_2\leftrightarrow R_3\\ -(a+1)R_2+R_3\to R_3}} \left(\begin{array}{ccc|c}1 & a & 1 & a\\ 0 & 1-a & a-1 & a^2-a\\ 0 & 0 & (1-a)(a+2) & (1+a)^2(1-a)\end{array}\right)$$

$$(a-1)(a+2)z=(1+a)^2(a-1)$$

Case 1: If $a\neq 1, a\neq -2$

$$z=\frac{(a+1)^2}{a+2}$$

$$(1-a)y+(a-1)z=a^2-a$$

$$a\neq 1, y=z-a=\frac{1}{a+2}$$

$$x=a-\frac{(a+1)^2-a}{a+2}=\frac{1}{a+2}-1$$

Case 2: If $a=-2, 0z=3$

Thus the system has no solutions.

Case 3: If $a=1$: $0z=0, z=t, t\in R$

$$(a-1)y+(a^2-1)t=a^3-1$$

$$0y=0, y=s, s\in R, \quad x=1-s-t$$

The solution is a plane, $\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}1\\0\\0\end{pmatrix}+s\begin{pmatrix}-1\\1\\0\end{pmatrix}+t\begin{pmatrix}-1\\0\\1\end{pmatrix}$.

Example 2: S2012 P13 Q10

Find the set of values of a for which the system of equations

$$\begin{aligned}x-2y-2z&=-7,\\ 2x+(a-9)y-10z&=-11,\\ 3x-6y+2az&=-29,\end{aligned}$$

has a unique solution. [4]

Show that the system has no solution in the case $a=-3$. [2]

Given that $a=5$,

(i) show that the number of solutions is infinite, [2]

(ii) find the solution for which $x+y+z=2$. [3]

$$\begin{vmatrix}1 & -2 & -2\\ 2 & a-9 & -10\\ 3 & -6 & 2a\end{vmatrix}\neq 0, \begin{vmatrix}a-9 & -10\\ -6 & 2a\end{vmatrix}+2\begin{vmatrix}2 & -10\\ 3 & 2a\end{vmatrix}-2\begin{vmatrix}2 & a-9\\ 3 & -6\end{vmatrix}\neq 0$$

$$2a(a-9)-60+2(4a+30)-2(-12-3a+27)\neq 0$$

$$a^2-9a-30+4a+30+12+3a-27\neq 0$$

$$a^2-2a-15\neq 0\Rightarrow (a-5)(a+3)\neq 0,$$

The set of values are $\{a: a\in R, a\neq 5, a\neq -3\}$.

$$x-2y-2z=-7-(1)$$

$$2x-12z-10z=-11-(2)$$

$$3x-6y-6z=-29-(3)$$

$(3)-3(1): 8=0$, hence the system has no solution.

(i)

$$x-2y-2z=-7-(1)$$

$$2x-4y-10z=-11-(2)$$

$$3x-6y+10z=-29-(3)$$

$$(2)+(3): 5x-10y=-40-(4), \quad (3)+5(1): 8x-16y=-64-(5)$$

Let $x=s, s\in R$. Thus $y=\frac{1}{2}(s+8)$, $z=\frac{1}{2}(s-s-8+7)=-\frac{1}{2}$,

hence the system has infinite solutions.

(ii) When $x+y+z=2, s+0.5s+4=2.5$

$$1.5s=-1.5\to s=-1$$

Thus $x=-1, y=3.5, z=-0.5$.

9.4 Linear Spaces

A) Linear Spaces

- LO: recall and use the axioms of a linear (vector) space (restricted to spaces of finite dimension over real numbers only);

Let *V* be an arbitrary nonempty set of objects which two operations are defined: addition, and scalar multiplication. By **addition** we mean a rule for associating each pair of objects **u** and **v** in *V* an object **u** + **v** , called **sum** of **u** and **v**; by **scalar multiplication** we mean a rule for associating each scalar *k* and each object **u** in *V* an object *k***u,** called **scalar multiple** of **u** by *k*. If the following axioms are satisfied by all objects **u**, **v**, **w** in *V* and all scalars *k* and *m*, then we call *V* a ***vector space*** or a ***linear space*** and we call the objects in *V* ***vectors***.

1. If **u** and **v** are objects in *V*, then **u** + **v** is in *V*. (closure under addition)
2. **u** + **v** = **v** + **u** (commutative under addition)
3. u + v + w = u + v + w (associative under addition)
4. There is an object **0** in *V*, called a zero vector for *V*, such that

   0 + **u** = **u** + 0 = **u** for all **u** in *V*. (identity element of addition)
5. For each **u** in *V*, there is an object in *V*, called a negative of **u**, such that

   u + (-u) = (-u) + u = 0. (inverse element)
6. If *k* is any scalar and u is any object in *V*, then *k***u** is in *V*. (closure under scalar multiplication)
7. $k(\boldsymbol{u}+\boldsymbol{v}) = k\boldsymbol{u} + k\boldsymbol{v}$ (distributive under scalar multiplication)
8. $(k+m)\boldsymbol{u} = k\boldsymbol{u} + m\boldsymbol{u}$(distributive under scalar multiplication)
9. $k(m\boldsymbol{u}) = (km)(\boldsymbol{u})$ (commutative under scalar multiplication)
10. 1**u** = **u** (identity element of scalar multiplication)

- Basically, a linear space must be closed, associative, and commutative under both operations. Usually in the exam we only verify closure and identity rules as most of the other rules hold for real numbers in general.

Examples of linear spaces

1. The set of all points in 3 dimensional space , R^3: (x, y, z)
2. The set of all points in 2 dimensional space (R^2) satisfying the equation $y = 2x$.
3. The set of all 2 by 2 matrices $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$.
4. The set of all degree 2 polynomials $ax^2 + bx + c$.
5. The set of all solutions to the differential equation $\frac{d^2y}{dx^2} - 3\frac{dy}{dx} + 2y = 0$.

Examples of sets that are not linear spaces

1. The set of all points in Quadrant I of 2 dimensional space, R^2: $(x, y), x \geq 0, y \geq 0$.
Reason: Rule 6 (multiply by negative scalars), Rule 5
2. The set of all points in 2 dimensional space satisfying the equation $y = 2x - 1$.
Reason: Rule 6 (Multiply by 0), Rule 4 ($(0, 0)$ is not in the space)
3. The set of all degree 2 polynomials $ax^2 + bx + c$, where $a \neq 0$.
Reason: Rule 4 (no zero vector), Rule 1 (add $-ax^2$ gives a vector outside the space)

B) Linear Independence

- LO: understand the idea of linear independence, and determine whether a given set of vectors is dependent or independent;
- LO: recall that the columns (or rows) of a square matrix are independent if and only if the determinant is non-zero.

Linear Combinations

Definition: If **w** is a vector in a vector space *V*, then **w** is said to be a ***linear combination*** of the vectors $v_1, v_2, v_3, \dots, v_r$ in *V* if **w** can be expressed in the form

$$\mathbf{w} = k_1v_1 + k_2v_2 + k_3v_3 + \cdots + k_rv_r$$

where $k_1, k_2, k_3, \dots, k_r$ are scalars. These scalars are called the ***coefficients*** of the linear combination.

Example: Let $v_1 = (2,1,-1),\ v_2 = (5,-1,0),\ v_3 = (-2,3,-1)$

1. $(11,-6,2) = v_1 + v_2 - 2v_3$ is a linear combination of the three vectors.
2. $(15,4,-5) = 5v_1 + v_2$ is a linear combination of the three vectors.
3. $(11,-6,1)$ is not a linear combination of the three vectors.

Proof: Solve the equation $k_1v_1 + k_2v_2 + k_3v_3 = (11,-6,1)$ for k_1, k_2 and k_3, which will not have a common solution.

Note: Exam questions usually features column vectors instead of row vectors. I used row vectors here just to save space.

Linear Independence

Definition 1: A set of vectors are ***linearly independent*** if and only if no vector in *S* is expressible as a linear combination of the other vectors in *S*.
If at least one of the vectors in *S* is expressible as a linear combination of the other vectors in *S* , then *S* is a ***linearly dependent*** set of vectors.

Definition 2: If $S = \{v_1, v_2, v_3, \dots, v_r\}$ is a nonempty set of vectors in a vector space *V*, then the vector equation

$$k_1v_1 + k_2v_2 + k_3v_3 + \cdots + k_rv_r = 0 (*)$$

has at least one solution, namely, $k_1 = k_2 = k_3 = \cdots = k_r = 0$. We call this the ***trivial solution*** (an obvious solution).
If this is the only solution, then *S* is said to be a ***linearly independent set***. If there are solutions in addition to the trivial solution, then *S* is said to be a ***linearly dependent set***.

- The above 2 definitions are equivalent. (*) can be re-arranged into

$$v_r = -\frac{k_1}{k_r}v_1 - \frac{k_2}{k_r}v_2 - \frac{k_3}{k_r}v_3 + \cdots + -\frac{k_{r-1}}{k_r}v_r$$

This form shows that v_r is always a linear combination of the other vectors unless $k_r = 0$. But if at least one of the other constants do not equal 0, at least one of the vectors in the set must be the zero vector and this makes the vectors linearly dependent. Thus the set of vectors must be linearly independent if and only if all of the constants $k_1, k_2, k_3, \dots, k_r$ equal 0.

Example:

1. The vectors $v_1 = (1,-2,3),\ v_2 = (5,6,-1),\ v_3 = (3,2,1)$ are linearly dependent as

$$v_3 = \frac{1}{2}v_1 + \frac{1}{2}v_2.$$

Bonus: If we form a matrix A using these three vectors as rows / columns, $\det(A)$ will equal 0 if the vectors are linearly dependent (if A is a square matrix).

2. The vectors $v_1 = (1,2,2,-1),\ v_2 = (4,9,9,-4),\ v_3 = (5,8,9,-5)$ are linearly independent. (Verify this by solving $k_1v_1 + k_2v_2 + k_3v_3 = 0$).

C) Subspaces

- LO: understand the idea of the subspace spanned by a given set of vectors;

Definition: A subset *W* of a vector space *V* is called a ***subspace*** of *V* if *W* is itself a vector space under the addition and scalar multiplication defined on *V*.

Generally, a subspace is a subset of the linear space that is itself a linear space.

Examples:

1. The set containing the zero vector is a subset of all linear spaces.

2. The set of all lines passing through the origin, $y = kx$ is a subspace of R^2.

3. The set of all degree 1 polynomials $ax + b$ is a subset of the set of degree 2 polynomials.

4. The set of all points in Quadrant I in R^2 is NOT a subspace of R^2. (as it is not a linear space, see the previous sections for explanation)

D) Bases and Dimensions

- LO: recall that a basis for a space is a linearly independent set of vectors that spans the space, and determine a basis in simple cases;
- LO: recall that the dimension of a space is the number of vectors in a basis;

Spanning

Definition: The subspace of a vector space *V* that is formed from all possible linear combinations of the vectors in a nonempty set *S* is called the ***span of S***, and we say that the vectors in $\boldsymbol{S}$ span that subspace. If $\boldsymbol{S} = \{w_1, w_2, \dots, w_r\}$, then we denote the span of *S* by

$$span(\boldsymbol{S}) \quad \text{or} \quad span(w_1, w_2, \dots, w_r)$$

Examples:

1. Determine if the vectors $(1,1,2), (1,0,1), (2,1,3)$ span R^3.
As $(1,1,2), +(1,0,1) = (2,1,3)$, the three vectors are linearly independent and cannot span R^3.
Alternatively, find a vector in R^3 that is not a linear combination of the three vectors, e.g. $(1,0,0)$.

2. The set of all points in 3 dimensional space , R^3: (x, y, z), $\boldsymbol{S}$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \rightarrow span(\boldsymbol{S}) = \left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}$$

3. The set of all degree 2 polynomials $ax^2 + bx + c$, *S*.

$$ax^2 + bx + c = a(x^2) + b(x) + c(1) \rightarrow span(\boldsymbol{S}) = \{1, x, x^2\}$$

4. The set of all solutions to the differential equation $\frac{d^2y}{dx^2} - 3\frac{dy}{dx} + 2y = 0$.

$y = Ae^x + Be^{2x}$, thus the set $\{e^x,\ e^{2x}\}$ spans the linear space.

Bases and Dimensions

Definition: If *V* is any vector space and $S = \{v_1, v_2, v_3, \dots, v_r\}$ is a finite set of vectors in *V*, then *S* is called a *basis* for *V* if the following two conditions hold:
(a) *S* is linearly independent. **(b)** *S* spans *V*.

Bonus: If $\boldsymbol{S} = \{v_1, v_2, v_3, \dots, v_r\}$is a basis for a vector space *V*, then every vector **v** in *V* can be expressed in the form $\mathbf{v} = k_1v_1 + k_2v_2 + k_3v_3 + \cdots + k_rv_r$ in exactly one way.
Basically, a basis for a space is a linearly independent set of vectors that spans the space.
Think of it as the basic cells that build up the body or the atoms that build up the world.

Definition: The ***dimension*** of a finite-dimensional vector space *V* is denoted by $\dim(V)$ and is defined to be the number of vectors in a basis for *V*. In addition, the zero vector space is defined to have dimension zero.
In applied mathematics, this is known as degrees of freedom (see Topic 17)
Bonus: The dimension of the space spanned by a linearly independent set of vectors is equal to the number of vectors in that set.

Examples: Let the linear space spanned by *S* from examples 3 – 4 be *V*.
5. From (2) above, $\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$ form a basis for R^3. R^3 has dimension 3 (hence 3-D).
6. From (3) above, $\{1, x, x^2\}$ is a basis of *V*. *V* has dimension 3.
7. From (4) above, $\{e^x,\ e^{2x}\}$ is a basis of *V*. *V* has dimension 2.

Example 1: S2002 P1 Q10

The vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ are given by

$$\mathbf{a}_1 = \begin{pmatrix}3\\2\\1\\0\end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \quad \mathbf{a}_3 = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \quad \mathbf{b}_1 = \begin{pmatrix}3\\2\\0\\2\end{pmatrix}, \quad \mathbf{b}_2 = \begin{pmatrix}2\\2\\0\\1\end{pmatrix}, \quad \mathbf{b}_3 = \begin{pmatrix}1\\1\\0\\1\end{pmatrix}.$$

The subspace of $\mathbb{R}^4$ spanned by $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ is denoted by V_1, and the subspace of $\mathbb{R}^4$ spanned by $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ is denoted by V_2. Show that V_1 and V_2 each have dimension 3. [3]

The set of vectors which belong to both V_1 and V_2 is denoted by V_3. Find a basis for V_3. [2]

The set of vectors which consists of the zero vector and all vectors which belong to only one of V_1 and V_2 is denoted by W.

(i) Write down two linearly independent vectors which belong to W. [2]

(ii) Show that W is not a linear space. [3]

$$c_1\begin{bmatrix}3\\2\\1\\0\end{bmatrix} + c_2\begin{bmatrix}1\\1\\0\\0\end{bmatrix} + c_3\begin{bmatrix}0\\0\\1\\0\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\end{bmatrix}$$

$3c_1 + c_2 = 0 - (1)$
$2c_1 + c_2 = 0 - (2)$
$(1) - (2)$: $c_1 = 0, c_2 = 0$
$c_1 + c_3 = 0 \rightarrow c_3 = 0$

$$c_1\begin{bmatrix}3\\2\\0\\2\end{bmatrix} + c_2\begin{bmatrix}2\\2\\0\\1\end{bmatrix} + c_3\begin{bmatrix}1\\1\\0\\1\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\end{bmatrix}$$

$3c_1 + 2c_2 + c_3 = 0 - (3)$
$2c_1 + 2c_2 + c_3 = 0 - (4)$
$2c_1 + c_2 + c_3 = 0 - (5)$
$(3) - (4)$: $c_1 = 0$, $(4) - (5)$: $c_2 = 0 \rightarrow c_3 = 0$

Since $(0, 0, 0)$ is the only solution to the systems above, vectors $\{a_1, a_2, a_3\}$ and vectors $\{b_1, b_2, b_3\}$ are linearly independent and are bases of V_1 and V_2 respectively.
Hence V_1 and V_2 both have dimension 3.

Vectors in V_1 are in the form $\begin{bmatrix}x\\y\\z\\0\end{bmatrix}$ and vectors in V_2 are in the form $\begin{bmatrix}x\\y\\0\\z\end{bmatrix}$.

Thus vectors in both spaces are in the form $\begin{bmatrix}x\\y\\0\\0\end{bmatrix}$ → a basis of V_3 is $\left\{\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\end{bmatrix}\right\}$.

(i) Two linearly independent vectors in W are $\begin{bmatrix}1\\1\\1\\0\end{bmatrix}$ and $\begin{bmatrix}1\\1\\0\\1\end{bmatrix}$.

(Note: many other vectors also satisfy this condition)

(ii) $\begin{bmatrix}1\\1\\1\\0\end{bmatrix}+\begin{bmatrix}1\\1\\0\\1\end{bmatrix}=\begin{bmatrix}2\\2\\1\\1\end{bmatrix}$, which is in neither V_1 nor V_2, hence it is also not in W.

As W is not closed under addition, it is not a vector space.

Example 2: W2007 P1 Q10

The vectors $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4$ are defined as follows:

$$\mathbf{b}_1=\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\quad \mathbf{b}_2=\begin{pmatrix}1\\1\\0\\0\end{pmatrix},\quad \mathbf{b}_3=\begin{pmatrix}1\\1\\1\\0\end{pmatrix},\quad \mathbf{b}_4=\begin{pmatrix}1\\1\\1\\1\end{pmatrix}.$$

The linear space spanned by $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ is denoted by V_1 and the linear space spanned by $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_4$ is denoted by V_2.

(i) Give a reason why $V_1 \cup V_2$ is not a linear space. [1]

(ii) State the dimension of the linear space $V_1 \cap V_2$ and write down a basis. [2]

Consider now the set V_3 of all vectors of the form $q\mathbf{b}_2 + r\mathbf{b}_3 + s\mathbf{b}_4$, where q, r, s are real numbers. Show that V_3 is a linear space, and show also that it has dimension 3. [3]

Determine whether each of the vectors

$$\begin{pmatrix}4\\4\\2\\5\end{pmatrix} \quad \text{and} \quad \begin{pmatrix}5\\4\\2\\5\end{pmatrix}$$

belongs to V_3 and justify your conclusions. [4]

(i) $b_1 + b_2 + b_3 = \begin{bmatrix}3\\2\\1\\0\end{bmatrix}$ is in V_1 and $b_1 + b_2 + b_4 = \begin{bmatrix}3\\2\\1\\1\end{bmatrix}$ is in V_2.

Since $\begin{bmatrix}3\\2\\1\\0\end{bmatrix}+\begin{bmatrix}3\\2\\1\\1\end{bmatrix}=\begin{bmatrix}6\\4\\2\\1\end{bmatrix}$ is in neither V_1 (in the form $\begin{bmatrix}a\\b\\c\\0\end{bmatrix}$) nor V_2 (in the form $\begin{bmatrix}a\\b\\c\\c\end{bmatrix}$),

it is not in $V_1 \cup V_2$. As $V_1 \cup V_2$ is not closed under addition, it is not a vector space.

(ii) Vectors in $V_1 \cap V_2$ are in the form $\begin{bmatrix}a\\b\\0\\0\end{bmatrix}$.

$V_1 \cap V_2$ has dimension 2 and a basis of it is $\{b_1, b_2\}$.

Vectors in V_3 are in the form $\begin{bmatrix}a\\a\\b\\c\end{bmatrix}$.We have $\begin{bmatrix}a\\a\\b\\c\end{bmatrix}+\begin{bmatrix}d\\d\\e\\f\end{bmatrix}=\begin{bmatrix}a+d\\a+d\\b+e\\c+f\end{bmatrix}$ and $m\begin{bmatrix}a\\a\\b\\c\end{bmatrix}=\begin{bmatrix}ma\\ma\\mb\\mc\end{bmatrix}$.

Since V_3 is closed under addition and scalar multiplication, V_3 is a linear space.

$$c_1\begin{bmatrix}1\\1\\0\\0\end{bmatrix}+c_2\begin{bmatrix}1\\1\\1\\0\end{bmatrix}+c_3\begin{bmatrix}1\\1\\1\\1\end{bmatrix}=\begin{bmatrix}0\\0\\0\\0\end{bmatrix} \quad \Big| \quad \begin{aligned}&c_1+c_2+c_3=0-(1)\\&c_2+c_3=0-(2)\\&c_3=0-(3)\end{aligned}$$

As $(0, 0, 0)$ is the only solution to the system, thus the vectors $\{b_2, b_3, b_4\}$ are linearly independent. V_3 has dimension 3.

As vectors in V_3 must be in the form $\begin{bmatrix} a \\ a \\ b \\ c \end{bmatrix}$, $\begin{bmatrix} 5 \\ 4 \\ 2 \\ 5 \end{bmatrix}$ does not belong to V_3.

$2b_2 - 3b_3 + 5b_4 = \begin{bmatrix} 4 \\ 4 \\ 2 \\ 5 \end{bmatrix}$, hence it belongs to V_3.

9.5 Linear Transformations

A) An Introduction to Linear Transformations

- LO: understand the use of matrices to represent linear transformations from $\mathbb{R}^n \to \mathbb{R}^m$;

Definition: If $T: V \to W$ is a function from a vector space *V* to a vector space *W*, then *T* is called a ***linear transformation*** from *V* to *W* if the following two properties hold for all vectors **u** and **v** in *V* and for all scalars *k*:

(i) $\boldsymbol{T(ku) = k\,T(u)}$ [Homogeneity property]

(ii) $\boldsymbol{T(u + v) = T(u) + T(v)}$ [Additive property]

In the special case where $V = W$, the linear transformation *T* is called a ***linear operator*** on the vector space *V*.

Every linear transformation can be represented by a matrix. For example;

1. Matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ represents a 90^o anti-clockwise rotation centred at the origin.

2. Matrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ represents a reflection along the straight line $y = x$.

Matrices can also be used to represent transformations from a dimension to another dimension, denoted by the notation $\mathbb{R}^n \to \mathbb{R}^m$. For example,

1. $\begin{bmatrix} 1 & 1 & 1 & 4 \\ 2 & 1 & 4 & 11 \\ 3 & 4 & 1 & 9 \\ 4 & -3 & 18 & 37 \end{bmatrix}$ represents a linear transformation $T: \mathbb{R}^4 \to \mathbb{R}^4$.

(A 4×1 column vector will still be a 4×1 column vector after the transformation)

2. $\begin{bmatrix} 1 & 3 & 5 & 7 & 9 \\ 2 & 4 & 6 & 8 & 10 \end{bmatrix}$ represents a linear transformation $T: \mathbb{R}^2 \to \mathbb{R}^5$.

(A 2×1 column vector will become a 5×1 column vector after the transformation)

3. $\begin{bmatrix} 1 & -5 & 2 \\ -1 & 3 & 8 \\ 2 & 0 & -4 \\ 3 & 1 & 0 \\ 4 & 0 & 0 \\ 0 & -2 & 1 \end{bmatrix}$ represents a linear transformation $T: \mathbb{R}^6 \to \mathbb{R}^3$.

(A 6×1 column vector will become a 3×1 column vector after the transformation)

B) The Subspaces of a Matrix

- LO: understand the terms ‘column space’, ‘row space’, range space’ and ‘null space’, and determine the dimensions of, and bases for, these spaces in simple cases;
- LO: determine the rank of a square matrix, and use (without proof) the relation between the rank, the dimension of the null space and the order of the matrix;

Four Subspaces of a Matrix

Definition: If A is an $m \times n$ matrix, then

1. The subspace of R^n spanned by the row vectors of A is called the ***row space*** of A,

2. The subspace of R^n spanned by the column vectors of A is called the ***column space*** or the ***range space*** of A.

3. The solution space of the homogeneous system of equations $Ax = 0$, which is a subspace of R^n, is called the ***null space*** of A.

Elementary row operations do not change the row space, column space, or null space of a matrix.

If a matrix R is in row echelon form, then the row vectors with the leading 1′s (the nonzero row vectors) form a basis for the row space of R, and the column vectors with the leading 1′s of the row vectors form a basis for the column space of R.

Example: Find the row space, column space and null space of the matrix

$$A = \begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 2 & -6 & 9 & -1 & 8 & 2 \\ 2 & -6 & 9 & -1 & 9 & 7 \\ -1 & 3 & -4 & 2 & -5 & -4 \end{bmatrix}$$

Solution: Performing ERO:

$$A = \begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 2 & -6 & 9 & -1 & 8 & 2 \\ 2 & -6 & 9 & -1 & 9 & 7 \\ -1 & 3 & -4 & 2 & -5 & -4 \end{bmatrix} \xrightarrow[-2R_1+R_3 \to R_3]{-2R_1+R_2 \to R_2} \xrightarrow{R_1+R_4 \to R_4} \begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 0 & 0 & 1 & 3 & -2 & -6 \\ 0 & 0 & 1 & 3 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{-R_2+R_3 \to R_3} \begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 0 & 0 & 1 & 3 & -2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

1. A basis for the row space of matrix A is

$$\{(1,-3,4,-2,5,4),(0,0,1,3,-2,-6),(0,0,0,0,1,5)\}$$

The row space of matrix A is

$$c_1(1,-3,4,-2,5,4) + c_2\,(0,0,1,3,-2,-6) + c_3\,(0,0,0,0,1,5)$$

OR: using the first three rows of the original matrix:

A basis for the row space of matrix A is

$$\{(1,-3,4,-2,5,4),(2,-6,9,-1,8,2),(2,-6,9,-1,9,7)\}$$

The row space of matrix A is

$$c_1(1,-3,4,-2,5,4) + c_2\,(2,-6,9,-1,8,2) + c_3\,(2,-6,9,-1,9,7)$$

Both answers a valid (the second one is a linear combination of the first one)

2. A basis for the column space of matrix A is $\left\{\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}4\\1\\0\\0\end{pmatrix}, \begin{pmatrix}5\\-2\\1\\0\end{pmatrix}\right\}$.

The column space of matrix A is $c_1\begin{pmatrix}1\\0\\0\\0\end{pmatrix} + c_2\begin{pmatrix}4\\1\\0\\0\end{pmatrix} + c_3\begin{pmatrix}5\\-2\\1\\0\end{pmatrix}$.

OR: using the first, third and fifth rows of the original matrix:

A basis for the column space of matrix A is $\left\{\begin{pmatrix}1\\2\\2\\-1\end{pmatrix}, \begin{pmatrix}4\\9\\9\\-4\end{pmatrix}, \begin{pmatrix}5\\8\\9\\-5\end{pmatrix}\right\}$.

The column space of matrix A is $c_1\begin{pmatrix}1\\2\\2\\-1\end{pmatrix} + c_2\begin{pmatrix}4\\9\\9\\-4\end{pmatrix} + c_3\begin{pmatrix}5\\8\\9\\-5\end{pmatrix}$.

3. $Ax = 0$

$$\begin{bmatrix}1 & -3 & 4 & -2 & 5 & 4\\ 0 & 0 & 1 & 3 & -2 & -6\\ 0 & 0 & 0 & 0 & 1 & 5\\ 0 & 0 & 0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\\x_5\\x_6\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\end{bmatrix}$$

3rd row: $x_5 + 5x_6 = 0$

Let $x_6 = t, t \in R, x_5 = -5t$

2nd row: $x_3 + 3x_4 - 2x_5 - 5x_6 = 0$

$x_3 + 3x_4 = -5t$

Let $x_4 = s, s \in R, x_3 = -3s - 5t$

1st row: $x_1 - 3x_2 + 4x_3 - 2x_4 + 5x_5 + 4x_6 = 0$

$x_1 - 3x_2 = 14s + 41t$

Let $x_2 = p, p \in R, x_1 = 3p + 14s + 41t$

The null space of matrix A is $\begin{bmatrix}3p + 14s + 41t\\ p\\ -3s - 5t\\ s\\ -5t\\ t\end{bmatrix} = p\begin{bmatrix}3\\1\\0\\0\\0\\0\end{bmatrix} + s\begin{bmatrix}14\\0\\-3\\1\\0\\0\end{bmatrix} + t\begin{bmatrix}41\\0\\-5\\0\\-5\\1\end{bmatrix}$.

A basis for the null space of matrix A is $\left\{\begin{bmatrix}3\\1\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}14\\0\\-3\\1\\0\\0\end{bmatrix}, \begin{bmatrix}41\\0\\-5\\0\\-5\\1\end{bmatrix}\right\}$.

Rank and Nullity

Definition: The row space and column space of a matrix *A* have the same dimension.
The common dimension of the row space (row rank) and column space (column rank) of a matrix *A* is called the ***rank*** of *A* and is denoted by rank (*A*);
The dimension of the null space of *A* is called the ***nullity*** of *A* and is denoted by nullity (*A*).

Example: From matrix A above,
1. The dimension of the range space of A / rank of A is 3.
2. The dimension of the null space of A / nullity of A is 3.

The Rank-Nullity Theorem: If *A* is a matrix with *n* columns, then

$$\mathbf{rank(A) + nullity(A) = \boldsymbol{n}}.$$

Proof: Since *A* has *n* columns, the homogeneous linear system has *n* unknowns (variables). These falls into two distinct categories: the leading variables and the free variables. Thus,

$$\mathbf{no.\ of\ leading\ variables + no.\ of\ free\ variables} = n$$

But the number of leading variables is the same as the number of leading 1′s in the reduced row echelon form of *A*, which is the rank of *A*; and the number of free variables is the same as the number of parameters in the general solution of $Ax = 0$, which is the nullity of *A*.

In the example above, $rank(A) = 3, nullity(A) = 3, n = 3 + 3 = 6$.

Example 1: S2008 P1 Q12 OR

The linear transformation $T : \mathbb{R}^4 \to \mathbb{R}^4$ is represented by the matrix

$$\begin{pmatrix} 1 & 2 & -1 & -1 \\ 1 & 3 & -1 & 0 \\ 1 & 0 & 3 & 1 \\ 0 & 3 & -4 & -1 \end{pmatrix}.$$

The range space of T is denoted by V.

(i) Determine the dimension of V. [3]

(ii) Show that the vectors $\begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 0 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 3 \\ -4 \end{pmatrix}$ are linearly independent. [4]

(iii) Write down a basis of V. [1]

The set of elements of $\mathbb{R}^4$ which do not belong to V is denoted by W.

(iv) State, with a reason, whether W is a vector space. [1]

(v) Show that if the vector $\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}$ belongs to W then $y - z - t \neq 0$. [5]

(i)

$$\mathbf{M} = \begin{bmatrix} 1 & 2 & -1 & -1 \\ 1 & 3 & -1 & 0 \\ 1 & 0 & 3 & 1 \\ 0 & 3 & -4 & -1 \end{bmatrix} \xrightarrow{\substack{-R_1+R_2 \to R_2 \\ -R_1+R_3 \to R_3}} \begin{bmatrix} 1 & 1 & 1 & 4 \\ 0 & 1 & 0 & 1 \\ 0 & -2 & 4 & 2 \\ 0 & 3 & -4 & -1 \end{bmatrix}$$

$$\xrightarrow{\substack{2R_2+R_3 \to R_3 \\ -3R_2+R_4 \to R_4}} \begin{bmatrix} 1 & 1 & 1 & 4 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 4 & 4 \\ 0 & 0 & -4 & -4 \end{bmatrix}$$

$$\xrightarrow{\substack{\frac{1}{4}R_3 \to R_3 \\ -4R_3+R_4 \to R_4}} \begin{bmatrix} 1 & 1 & 1 & 4 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Hence $\dim(V) = 3$.

(ii)

$$c_1 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 3 \\ 0 \\ 3 \end{bmatrix} + c_3 \begin{bmatrix} -1 \\ -1 \\ 3 \\ -4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$c_1 + 2c_2 - c_3 = 0 - (1) \qquad c_1 + 3c_2 - c_3 = 0 - (2)$$

$(2) - (1)$: $c_2 = 0$. Since $3c_2 - 4c_3 = 0, c_3 = 0$. $c_1 = c_3 - 2c_2 = 0$.

$(0, 0, 0)$ is the only solution to the system, thus the vectors are linearly independent.

(iii)

A basis of V is 3 linearly independent vectors in V, for example $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 3 \\ -4 \end{bmatrix} \right\}$.

(iv)

Since $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ is in V, $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ is not in W. As W does not contain the zero vector, it is not closed under scalar multiplication. Thus W is not a vector space.

(v)

Vectors in V have the form $\mathbf{v} = c_1\begin{bmatrix}1\\1\\1\\0\end{bmatrix} + c_2\begin{bmatrix}2\\3\\0\\3\end{bmatrix} + c_3\begin{bmatrix}-1\\-1\\3\\-4\end{bmatrix}$.

When $x = c_1 + 2c_2 - c_3, y = c_1 + 3c_2 - c_3, z = c_1 + 3c_3,$

$$y - z = 3c_2 - 4c_3, \text{thus } \mathbf{v} = \begin{bmatrix}x\\y\\z\\y-z\end{bmatrix}$$

If $y - z - t = 0, \mathbf{v} = \begin{bmatrix}x\\y\\z\\t\end{bmatrix}$, thus **v** is in V, **v** is not in W.

Contrapositive of the statement: If **v** is in W, then $y - z - t \neq 0$.

Note: Statements (if p then q) and (if not q then not p) are identical.

Example 2: S2012 P11 Q7

The linear transformations $T_1 : \mathbb{R}^4 \to \mathbb{R}^4$ and $T_2 : \mathbb{R}^4 \to \mathbb{R}^4$ are represented by the matrices

$$\mathbf{M}_1 = \begin{pmatrix}1 & 1 & 1 & 4\\2 & 1 & 4 & 11\\3 & 4 & 1 & 9\\4 & -3 & 18 & 37\end{pmatrix} \quad \text{and} \quad \mathbf{M}_2 = \begin{pmatrix}1 & 1 & 1 & -1\\2 & 3 & 0 & 1\\3 & 4 & 1 & 0\\4 & 5 & 2 & 0\end{pmatrix},$$

respectively. The null space of T_1 is denoted by K_1 and the null space of T_2 is denoted by K_2. Show that the dimension of K_1 is 2 and that the dimension of K_2 is 1. [5]

Find the basis of K_1 which has the form $\left\{\begin{pmatrix}p\\q\\1\\0\end{pmatrix}, \begin{pmatrix}r\\s\\0\\1\end{pmatrix}\right\}$ and show that K_2 is a subspace of K_1. [5]

$$\mathbf{M}_1 = \begin{bmatrix}1 & 1 & 1 & 4\\2 & 1 & 4 & 11\\3 & 4 & 1 & 9\\4 & -3 & 18 & 37\end{bmatrix} \xrightarrow[-4R_1+R_4\to R_4]{\substack{-2R_1+R_2\to R_2\\-3R_1+R_3\to R_3}} \begin{bmatrix}1 & 1 & 1 & 4\\0 & -1 & 2 & 3\\0 & 1 & -2 & -3\\0 & -7 & 14 & 21\end{bmatrix} \xrightarrow[3R_2+R_4\to R_4]{\substack{-R_2\to R_2\\-R_2+R_3\to R_3}} \begin{bmatrix}1 & 1 & 1 & 4\\0 & 1 & -2 & -3\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{bmatrix}$$

$$\mathbf{M}_2 = \begin{bmatrix}1 & 1 & 1 & -1\\2 & 3 & 0 & 1\\3 & 4 & 1 & 0\\4 & 5 & 2 & 0\end{bmatrix} \xrightarrow[-4R_1+R_4\to R_4]{\substack{-2R_1+R_2\to R_2\\-3R_1+R_3\to R_3}} \begin{bmatrix}1 & 1 & 1 & 4\\0 & 1 & -2 & 3\\0 & 1 & -2 & 3\\0 & 1 & -2 & 4\end{bmatrix} \xrightarrow[-R_2+R_4\to R_4]{\substack{R_3\leftrightarrow R_4\\-R_2+R_3\to R_3}} \begin{bmatrix}1 & 1 & 1 & 4\\0 & 1 & -2 & 3\\0 & 0 & 0 & 1\\0 & 0 & 0 & 0\end{bmatrix}$$

$rank\ (M_1) = 2$

$nullity = 4 - 2 = 2$

$$\begin{bmatrix}1 & 1 & 1 & 4\\0 & 1 & -2 & -3\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\end{bmatrix}$$

$$x_2 - 2x_3 - 3x_4 = 0$$

Let $x_3 = s, x_4 = t\ , s, t \in R,$

$rank(M_2) = 3$

$nullity = 4 - 3 = 1$

$$\begin{bmatrix}1 & 1 & 1 & 4\\0 & 1 & -2 & 3\\0 & 0 & 0 & 1\\0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\end{bmatrix}$$

$$x_4 = 0$$

$$x_2 - 2x_3 - 3x_4 = 0$$

thus $x_2 = 2s + 3t$.

$x_1 + x_2 + x_3 + 4x_4 = 0$

$x_1 = -2s - 3t - s - 4t = -3s - 7t$

Null space of K_1 is $\left\{\begin{bmatrix} -3s-7t \\ 2s+3t \\ s \\ t \end{bmatrix}\right\} \rightarrow$

A basis of K_1 is $\left\{\begin{bmatrix} -3 \\ 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -7 \\ 3 \\ 0 \\ 1 \end{bmatrix}\right\}$.

Let $x_3 = s, s \in R$, thus $x_2 = 2s$.

$x_1 + x_2 + x_3 + 4x_4 = 0$

$x_1 = -2s - s = -3s$

Null space of K_2 is $\left\{\begin{bmatrix} -3s \\ 2s \\ s \\ 0 \end{bmatrix}\right\} \rightarrow$

A basis of K_2 is $\left\{\begin{bmatrix} -3 \\ 2 \\ 1 \\ 0 \end{bmatrix}\right\}$.

Since all elements of K_2 are in K_1, K_2 is a subspace of K_1.

Example 3: S2006 P1 Q4

The linear transformation T : $\mathbb{R}^4 \rightarrow \mathbb{R}^4$ is represented by the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 2 & 3 \\ 2 & -3 & 4 & 5 \\ 5 & -6 & 10 & 14 \\ 4 & -5 & 8 & 11 \end{pmatrix}.$$

Show that the dimension of the range space of T is 2. [3]

Let **M** be a given 4×4 matrix and let S be the vector space consisting of vectors of the form **MAx**, where $\mathbf{x} \in \mathbb{R}^4$. Show that if **M** is non-singular then the dimension of S is 2. [4]

$$\mathbf{M}_1 = \begin{bmatrix} 1 & -1 & 2 & 3 \\ 2 & -3 & 4 & 5 \\ 5 & -6 & 10 & 14 \\ 4 & -5 & 8 & 11 \end{bmatrix} \xrightarrow[-5R_1+R_3 \rightarrow R_3]{-2R_1+R_2 \rightarrow R_2} \begin{bmatrix} 1 & -1 & 2 & 3 \\ 0 & -1 & 0 & -1 \\ 0 & -1 & 0 & -1 \\ 0 & -1 & 0 & -1 \end{bmatrix}$$

$$\xrightarrow[-R_2+R_3 \rightarrow R_3]{-R_2 \rightarrow R_2} \begin{bmatrix} 1 & -1 & 2 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus $rank\ (A) = 2$.

Vectors in T have the form $\mathbf{x} = pb_1 + qb_2$, where $p, q \in R$ and $\{b_1, b_2\}$ is a basis of T.

Thus vectors in S have the form $\mathbf{x} = p\mathbf{M}b_1 + q\mathbf{M}b_2$.

If $\mathbf{M}b_1$ and $\mathbf{M}b_2$ are linearly dependent, $\mathbf{M}b_2 = k\mathbf{M}b_1$ for some $k \in R$.

Since $\mathbf{M}^{-1}$exists, $\mathbf{M}^{-1}\mathbf{M}b_2 = \mathbf{M}^{-1}k\mathbf{M}b_1 \rightarrow \mathbf{I}b_2 = k\mathbf{M}^{-1}\mathbf{M}b_1 \rightarrow b_2 = k\mathbf{I}b_1 = kb_1$.

Thus b_1 and b_2 are linearly dependent. Hence if b_1 and b_2 are linearly independent, so are $\mathbf{M}b_1$ and $\mathbf{M}b_2$.

$\{\mathbf{M}b_1, \mathbf{M}b_2\}$ is a basis of S $\rightarrow \dim(S) = 2$.

Example 4: S2009 P1 Q12 OR

The linear transformations $T_1 : \mathbb{R}^4 \to \mathbb{R}^4$ and $T_2 : \mathbb{R}^4 \to \mathbb{R}^4$ are represented by the matrices $\mathbf{M}_1$ and $\mathbf{M}_2$, respectively, where

$$\mathbf{M}_1 = \begin{pmatrix} 1 & 1 & 1 & 2 \\ 1 & 4 & 7 & 8 \\ 1 & 7 & 11 & 13 \\ 1 & 2 & 5 & 5 \end{pmatrix}, \qquad \mathbf{M}_2 = \begin{pmatrix} 2 & 0 & -1 & -1 \\ 5 & 1 & -3 & -3 \\ 3 & -1 & -1 & -1 \\ 13 & -1 & -6 & -6 \end{pmatrix}.$$

(i) Find a basis for R_1, the range space of T_1. [4]

(ii) Find a basis for K_2, the null space of T_2, and hence show that K_2 is a subspace of R_1. [5]

The set of vectors which belong to R_1 but do not belong to K_2 is denoted by W.

(iii) State whether W is a vector space, justifying your answer. [1]

The linear transformation $T_3 : \mathbb{R}^4 \to \mathbb{R}^4$ is the result of applying T_1 and then T_2, in that order.

(iv) Find the dimension of the null space of T_3. [3]

(i)

$$\mathbf{M}_1 = \begin{bmatrix} 1 & 1 & 1 & 2 \\ 1 & 4 & 7 & 8 \\ 1 & 7 & 11 & 13 \\ 1 & 2 & 5 & 5 \end{bmatrix} \xrightarrow[-R_1+R_4\to R_4]{\substack{-R_1+R_2\to R_2 \\ -R_1+R_3\to R_3}} \begin{bmatrix} 1 & 1 & 1 & 2 \\ 0 & 3 & 6 & 6 \\ 0 & 6 & 10 & 11 \\ 0 & 1 & 4 & 3 \end{bmatrix} \xrightarrow[-2R_2+R_3\to R_3]{\substack{\frac{1}{3}R_2\to R_2 \\ -R_2+R_4\to R_4}} \begin{bmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & -2 & -1 \\ 0 & 0 & 2 & 1 \end{bmatrix} \xrightarrow[0.5R_3\to R_3]{\substack{R_3+R_4\to R_4 \\ R_3\leftrightarrow R_4}} \begin{bmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & 0.5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$\text{rank}(\mathbf{M}_1) = 3$. A basis for R_1 is $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ 7 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 7 \\ 11 \\ 5 \end{bmatrix} \right\}$.

(ii)

$$\mathbf{M}_2 = \begin{bmatrix} 2 & 0 & -1 & -1 \\ 5 & 1 & -3 & -3 \\ 3 & -1 & -1 & -1 \\ 13 & -1 & -6 & -6 \end{bmatrix} \xrightarrow[-6R_1+R_4\to R_4]{\substack{-3R_1+R_2\to R_2 \\ -R_1+R_3\to R_3}} \begin{bmatrix} 2 & 0 & -1 & -1 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 \end{bmatrix} \xrightarrow[-R_2\to R_2]{\substack{R_2+R_3\to R_3 \\ R_2+R_4\to R_4}} \begin{bmatrix} 2 & 0 & -1 & -1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 & -1 & -1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$x_1 - x_2 = 0$. Let $x_1 = s, s \in R$, thus $x_2 = s$.

$2x_1 = x_3 + x_4$

Let $x_3 = t, t \in R$, thus $x_4 = 2s - t$

Null space of K_2 is $\left\{ \begin{bmatrix} s \\ s \\ t \\ 2s - t \end{bmatrix} \right\} \to$

A basis of K_2 is $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix} \right\}$.

$$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix} = 1.5 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 4 \\ 7 \\ 2 \end{bmatrix} + 0.5 \begin{bmatrix} 1 \\ 7 \\ 11 \\ 5 \end{bmatrix} \qquad \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix} = -0.5 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 4 \\ 7 \\ 2 \end{bmatrix} - 0.5 \begin{bmatrix} 1 \\ 7 \\ 11 \\ 5 \end{bmatrix}$$

As every basis vector in K_2 is a linear combination of the basis vectors of R_1, K_2 is a subspace of R_1.

(iii) Since $\begin{bmatrix}0\\0\\0\\0\end{bmatrix}$ is in both R_1 and K_2, $\begin{bmatrix}0\\0\\0\\0\end{bmatrix}$ is not in W. As W does not contain the zero vector, it is not closed under scalar multiplication. Thus W is not a vector space.

(iv) Vectors in R_1 are in the form $\mathbf{M}_1\mathbf{x} = ab_1 + bb_2 + cb_3$, where $a, b, c \in R$ and $\{b_1, b_2, b_3\}$ is its basis, 2 of which are a basis of K_2.

$$\begin{aligned}\mathbf{M}_3\mathbf{x} &= \mathbf{M}_2\mathbf{M}_1\mathbf{x}\\ &= \mathbf{M}_2(ab_1 + bb_2 + cb_3)\\ &= a\mathbf{M}_2b_1 + b\mathbf{M}_2b_2 + c\mathbf{M}_2b_3\end{aligned}$$

as vectors $\mathbf{M}_2b_1, \mathbf{M}_2b_2, \mathbf{M}_2b_3$ are linearly independent, $\dim(\mathbf{M}_3) = 3$.
Dimension of null space of $\mathbf{M}_3 = 4 - 3 = 1$.

C) Solving Systems of Linear Equations

- LO: use methods associated with linear spaces in the context of the solution of a set of linear equations;

To solve a system of linear equations in the form of $Ax = b$:

If A^{-1} is defined

$\text{nullity}(A) = 0$,

$x = A^{-1}b$.

If A^{-1} is undefined, $\text{nullity}(A) \neq 0$:

Step 1: find the null space of A.

Step 2: find a vector x_0 such that $Ax_0 = b$.

The solution is $x = x_0 +$ null space of A.

This is analogous to the method used to solve differential equations in Topic 6, with the null space being similar the complementary function and x_0 being similar to the particular integral.

Example 1: Given matrix $A = \begin{bmatrix}1 & -3 & 4 & -2 & 5 & 4\\ 2 & -6 & 9 & -1 & 8 & 2\\ 2 & -6 & 9 & -1 & 9 & 7\\ -1 & 3 & -4 & 2 & -5 & -4\end{bmatrix}$, solve the system $Ax = \begin{bmatrix}1\\2\\2\\-1\end{bmatrix}$.

From the previous example, the null space of A is $p\begin{bmatrix}3\\1\\0\\0\\0\\0\end{bmatrix} + s\begin{bmatrix}14\\0\\-3\\1\\0\\0\end{bmatrix} + t\begin{bmatrix}41\\0\\-5\\0\\-5\\1\end{bmatrix}$.

Since $A\begin{bmatrix}1\\0\\0\\0\\0\\0\end{bmatrix} = \begin{bmatrix}1\\2\\2\\-1\end{bmatrix}$, the solution is $x = \begin{bmatrix}1\\0\\0\\0\\0\\0\end{bmatrix} + p\begin{bmatrix}3\\1\\0\\0\\0\\0\end{bmatrix} + s\begin{bmatrix}14\\0\\-3\\1\\0\\0\end{bmatrix} + t\begin{bmatrix}41\\0\\-5\\0\\-5\\1\end{bmatrix}$.

Example 2: W2002 P1 Q10

The linear transformation $T : \mathbb{R}^4 \to \mathbb{R}^4$ is represented by the matrix $\mathbf{H}$, where

$$\mathbf{H} = \begin{pmatrix} 1 & 2 & -3 & -5 \\ -1 & 4 & 5 & 1 \\ 2 & 3 & 0 & -3 \\ -3 & 5 & 7 & 2 \end{pmatrix}.$$

(i) Find the dimension of the range space of T. [3]

(ii) Find a basis for the null space of T. [3]

(iii) It is given that $\mathbf{x}$ satisfies the equation

$$\mathbf{Hx} = \begin{pmatrix} 2 \\ -10 \\ -1 \\ -15 \end{pmatrix}.$$

Using the fact that

$$\mathbf{H}\begin{pmatrix} 1 \\ -3 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \\ -10 \\ -1 \\ -15 \end{pmatrix},$$

find the least possible value of $|\mathbf{x}|$. [7]

[For the vector $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$, $|\mathbf{x}| = \sqrt{(x_1^2 + x_2^2 + x_3^2 + x_4^2)}$.]

(i)

$$\mathbf{H} = \begin{bmatrix} 1 & 2 & -3 & -5 \\ -1 & 4 & 5 & 1 \\ 2 & 3 & 0 & -3 \\ -3 & 5 & 7 & 2 \end{bmatrix} \xrightarrow[-2R_1+R_3 \to R_3]{R_1+R_2 \to R_2 \\ 3R_1+R_4 \to R_4} \begin{bmatrix} 1 & 2 & -3 & -5 \\ 0 & 6 & 2 & -4 \\ 0 & -1 & 6 & 7 \\ 0 & 11 & -2 & -13 \end{bmatrix}$$

$$\xrightarrow[6R_2+R_3 \to R_3 \\ 11R_2+R_4 \to R_4]{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 2 & -3 & -5 \\ 0 & -1 & 6 & 7 \\ 0 & 0 & 38 & 38 \\ 0 & 0 & 64 & 64 \end{bmatrix}$$

$$\xrightarrow[\frac{1}{38}R_3 \to R_3 \\ -R_3+R_4 \to R_4]{-R_2 \to R_2} \begin{bmatrix} 1 & 2 & -3 & -5 \\ 0 & 1 & -6 & -7 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus the range space of T has dimension 3.

(ii)

$$\begin{bmatrix} 1 & 2 & -3 & -5 \\ 0 & 1 & -6 & -7 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$x_3 + x_4 = 0$

Let $x_4 = s,\ s \in R$, thus $x_3 = -s$.

$x_2 = 6x_3 + 7x_4 = s$

$x_1 + 2x_2 - 3x_3 - 5x_4 = 0$

$x_1 = 5s - 3s - 2s = 0$

Null space of T is $\left\{\begin{bmatrix} 0 \\ s \\ -s \\ s \end{bmatrix}\right\} \to$

A basis of T is $\left\{\begin{bmatrix} 0 \\ 1 \\ -1 \\ 1 \end{bmatrix},\right\}$.

(iii)

Since the null space of T is $s\begin{bmatrix} 0 \\ 1 \\ -1 \\ 1 \end{bmatrix}$, $\mathbf{H}s\begin{bmatrix} 0 \\ 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} - (1)$

$$\mathbf{H}\begin{bmatrix}1\\-3\\1\\-2\end{bmatrix}=\begin{bmatrix}2\\-10\\-1\\15\end{bmatrix}-(2),\ (1)+(2):\ \mathbf{H}\left\{\begin{bmatrix}1\\-3\\1\\-2\end{bmatrix}+s\begin{bmatrix}0\\1\\-1\\1\end{bmatrix}\right\}=\begin{bmatrix}2\\-10\\-1\\15\end{bmatrix}$$

Thus solution of $\mathbf{Hx}=\begin{bmatrix}2\\-10\\-1\\15\end{bmatrix}$ is $\mathbf{x}=\begin{bmatrix}1\\-3+s\\1-s\\-2+s\end{bmatrix}$.

$$|\mathbf{x}|^2=1^2+(s-3)^2+(1-s)^2+(s-2)^2=3s^2-12s+15$$

$$\frac{d|\mathbf{x}|^2}{ds}=6s-12,\qquad \frac{d^2|\mathbf{x}|^2}{ds^2}=6,$$

thus $|\mathbf{x}|^2$ is minimum when $6s-12=0\rightarrow s=2$.

$$|\mathbf{x}|_{min}=\sqrt{3(4)-24+15}=\sqrt{3}$$

Example 3: S2007 P1 Q12 OR

The linear transformation T : $\mathbb{R}^4\rightarrow\mathbb{R}^4$ is represented by the matrix

$$\mathbf{M}=\begin{pmatrix}1&-2&2&4\\2&-4&5&9\\3&-6&8&14\\5&-10&12&22\end{pmatrix}.$$

(i) Find the rank of **M**. [3]

(ii) Obtain a basis for the null space, K, of T. [3]

(iii) Evaluate $\mathbf{M}\begin{pmatrix}-1\\2\\-3\\4\end{pmatrix}$, and hence show that any solution of $\mathbf{Mx}=\begin{pmatrix}5\\11\\17\\27\end{pmatrix}$ $(*)$ has the form $\begin{pmatrix}-1\\2\\-3\\4\end{pmatrix}+\lambda\mathbf{e}_1+\mu\mathbf{e}_2$, where λ and μ are constants and $\{\mathbf{e}_1, \mathbf{e}_2\}$ is a basis for K. [3]

(iv) Find the solution $\mathbf{x}_1$ of $(*)$ such that the first component of $\mathbf{x}_1$ is A, and the sum of all the components of $\mathbf{x}_1$ is B. [5]

(i)

$$\mathbf{M}=\begin{bmatrix}1&-2&2&4\\2&-4&5&9\\3&-6&8&14\\5&-10&12&22\end{bmatrix}\xrightarrow[-5R_1+R_4\rightarrow R_4]{-2R_1+R_2\rightarrow R_2,\ -3R_1+R_3\rightarrow R_3}\begin{bmatrix}1&-2&2&4\\0&0&1&1\\0&0&2&2\\0&0&2&2\end{bmatrix}$$

$$\xrightarrow[-2R_2+R_4\rightarrow R_4]{-2R_2+R_3\rightarrow R_3}\begin{bmatrix}1&-2&2&4\\0&0&1&1\\0&0&0&0\\0&0&0&0\end{bmatrix}$$

$rank(\mathbf{M})=2.$

(ii)

$$\begin{bmatrix}1&-2&2&4\\0&0&1&1\\0&0&0&0\\0&0&0&0\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=\begin{bmatrix}0\\0\\0\\0\end{bmatrix}$$

$$x_3+x_4=0$$

Let $x_4=-s,\ s\in R$, thus $x_3=s$.

$x_1 - 2x_2 + 2x_3 + 4x_4 = 0$

Let $x_2 = t,\ t \in R$,

thus $x_1 = 2t - 2s + 4s = 2t + 2s$.

Null space of K is $\left\{\begin{bmatrix} 2s+2t \\ s \\ t \\ -t \end{bmatrix}\right\} \rightarrow$

A basis of K is $\left\{\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \\ -1 \end{bmatrix}\right\}$.

(iii)

$$\mathbf{M}\begin{bmatrix} -1 \\ 2 \\ -3 \\ 4 \end{bmatrix} = \begin{bmatrix} -1-4-6+16 \\ -2-8-15+36 \\ -3-12-24+56 \\ -5-20-36+88 \end{bmatrix} = \begin{bmatrix} 5 \\ 11 \\ 17 \\ 27 \end{bmatrix} - (1)$$

$\mathbf{e_1} = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{e_2} = \begin{bmatrix} 2 \\ 0 \\ 1 \\ -1 \end{bmatrix}$.

Since null space of K is $\lambda\mathbf{e_1} + \mu\mathbf{e_2}$, where $\lambda, \mu \in R,\ \mathbf{M}(\lambda\mathbf{e_1} + \mu\mathbf{e_2}) = 0 - (2)$.

$$(1) + (2): \mathbf{M}\left(\begin{bmatrix} -1 \\ 2 \\ -3 \\ 4 \end{bmatrix} + \lambda\mathbf{e_1} + \mu\mathbf{e_2}\right) = \begin{bmatrix} 5 \\ 11 \\ 17 \\ 27 \end{bmatrix}$$

Thus $\mathbf{x} = \begin{bmatrix} -1 \\ 2 \\ -3 \\ 4 \end{bmatrix} + \lambda\mathbf{e_1} + \mu\mathbf{e_2}$ is the solution to $\mathbf{Mx} = \begin{bmatrix} 5 \\ 11 \\ 17 \\ 27 \end{bmatrix}$.

(iv)

$$\begin{bmatrix} -1 \\ 2 \\ -3 \\ 4 \end{bmatrix} + \lambda\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \mu\begin{bmatrix} 2 \\ 0 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} A \\ p \\ q \\ B - A - p - q \end{bmatrix}$$

3rd and 4th rows:

$$-3 + \mu = q - (3)$$

$$4 - \mu = B - A - p - q - (4)$$

$$(3) + (4): B - A - p = 1 \rightarrow p = B - A - 1$$

(2), (3) into (1):

$$-1 + 2(p-2) + 2(q+3) = A$$

$$2q = A - 2p - 1$$

$$q = \frac{1}{2}(3A - 2B + 1)$$

Thus $\mathbf{x} = \begin{bmatrix} A \\ B - A - 1 \\ 0.5 - B + 1.5A \\ 0.5 - 1.5A + B \end{bmatrix}$.

9.6 Eigenvalues and Eigenvectors

- Linear algebra focuses on solving two types of equations:

(i) The equation $\mathbf{Ax} = \mathbf{b}$, which is what we investigated in previous sections: (ii) The equation $\mathbf{Ax} = \lambda\mathbf{x}$, which is what we will investigate here.

A) Eigenvalues and Eigenvectors

- LO: understand the terms 'eigenvalue' and 'eigenvector', as applied to square matrices;
- LO: find eigenvalues and eigenvectors of 2×2 and 3×3 matrices (restricted to cases where the eigenvalues are real and distinct);

- In the equation $\mathbf{Ax} = \lambda\mathbf{x}$, λ is known as the eigenvalue of the matrix **A** with **x** as its corresponding eigenvector.
- This equation literally describes vectors **x** which are only scaled when they are transformed using the matrix **A**.
- We will solve this system by introducing the identity matrix **I**, since when we factor out matrix **x** the difference between a matrix and a scalar cannot be evaluated.

$$\mathbf{Ax} - \boldsymbol{\lambda}\mathbf{x} = \mathbf{0}$$
$$\mathbf{Ax} - \boldsymbol{\lambda}\mathbf{Ix} = \mathbf{0}$$
$$\mathbf{x}(\mathbf{A} - \boldsymbol{\lambda}\mathbf{I}) = \mathbf{0}$$
$$\mathbf{x} = (\mathbf{A} - \boldsymbol{\lambda}\mathbf{I})^{-1}\mathbf{0} = \mathbf{0}$$

- If $(\mathbf{A} - \boldsymbol{\lambda}\mathbf{I})^{-1}$ exists, vector **x** must be the zero vector, hence infinite eigenvalues exist to this problem and this is not the case we are interested in.
- If $(\mathbf{A} - \boldsymbol{\lambda}\mathbf{I})^{-1}$does not exist, then we will only have specific solutions to λ and **x**, and hence this is the case we are interested in.
- Hence to find the eigenvalues of a matrix, $(\mathbf{A} - \boldsymbol{\lambda}\mathbf{I})^{-1}$ cannot exist, thus

The eigenvalues λ of a matrix **A** are obtained by solving the equation $\det(\mathbf{A} - \lambda\mathbf{I}) = 0.$ This is known as the characteristic equation / polynomial of the matrix.

- The eigenvectors to each eigenvalue are then obtained by solving $\mathbf{Ax} = \lambda\mathbf{x}$ for each eigenvalue found. Note that any scalar multiple of the eigenvectors found are equally valid answers.
- If an eigenvalue of a square matrix **A** is λ and **x** is a corresponding eigenvector, and if k is any positive integer, then is an eigenvalue of $\mathbf{A}^k$ is $\lambda^k$and **x** is a corresponding eigenvector. This can be proved using mathematical induction, but specific cases of it are shown and used in the examples below.

Example 1: W2008 P1 Q4

- We will first start with a definition question.

The matrix **A** has λ as an eigenvalue with **e** as a corresponding eigenvector. Show that **e** is an eigenvector of $\mathbf{A}^2$ and state the corresponding eigenvalue. [3]

Given that one eigenvalue of **A** is 3, find an eigenvalue of the matrix $\mathbf{A}^4 + 3\mathbf{A}^2 + 2\mathbf{I}$, justifying your answer. [3]

$$\mathbf{Ae} = \lambda\mathbf{e}$$
$$\mathbf{A}^2\mathbf{e} = \mathbf{A}(\mathbf{Ae}) = \mathbf{A}(\lambda\mathbf{e}) = \lambda(\mathbf{Ae})$$
$$= \lambda(\lambda\mathbf{e}) = \lambda^2\mathbf{e}$$

Thus matrix $\mathbf{A}^2$has eigenvector $\mathbf{e}$ with corresponding eigenvalue λ^2.

$$\begin{aligned}(\mathbf{A}^4 + 3\mathbf{A}^2 + 2\mathbf{I})\mathbf{e} &= \mathbf{A}^4\mathbf{e} + 3\mathbf{A}^2\mathbf{e} + 2\mathbf{I}\mathbf{e}\\ &= \mathbf{A}^2(\mathbf{A}^2\mathbf{e}) + 3(\lambda^2\mathbf{e}) + 2\mathbf{e}\\ &= \mathbf{A}^2(\lambda^2\mathbf{e}) + 3\lambda^2\mathbf{e} + 2\mathbf{e}\\ &= \lambda^2(\mathbf{A}^2\mathbf{e}) + 3\lambda^2\mathbf{e} + 2\mathbf{e}\\ &= \lambda^2(\lambda^2\mathbf{e}) + 3\lambda^2\mathbf{e} + 2\mathbf{e} = (\lambda^4 + 3\lambda^2 + 2)\mathbf{e}\end{aligned}$$

Matrix $\mathbf{A}^4 + 3\mathbf{A}^2 + 2\mathbf{I}$ has eigenvector $\mathbf{e}$ with corresponding eigenvalue $\lambda^4 + 3\lambda^2 + 2$.

$\lambda = 3$: Eigenvalue is $3^4 + 3(3^2) + 2 = 110$.

Example 2: S2002 P1 Q1

- For simplicity this one deals with a 2×2 matrix, and the next one deals with the most common case in examinations, 3×3 matrices.

Find the eigenvalues of the matrix

$$\begin{pmatrix} 4 & -2 \\ 3 & -1 \end{pmatrix}.$$

Find also corresponding eigenvectors. [5]

Eigenvalues:

$$\begin{vmatrix} 4-\lambda & -2 \\ 3 & -1-\lambda \end{vmatrix} = 0$$

$$(4-\lambda)(-1-\lambda)+6=0$$

$$\lambda^2 - 3\lambda + 2 = 0$$

$$(\lambda-1)(\lambda-2)=0,\ \lambda = 1 \text{ or } \lambda = 2$$

When $\lambda = 1$:	When $\lambda = 2$:
$\begin{bmatrix} 4 & -2 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 1\begin{bmatrix} x \\ y \end{bmatrix}$	$\begin{bmatrix} 4 & -2 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 2\begin{bmatrix} x \\ y \end{bmatrix}$
$4x - 2y = x,\ \ 3x - y = y$	$4x - 2y = 2x,\ \ 3x - y = 2y$
$2y = 3x$	$2y = 2x,\ \ \ \ 3x = 3y$
Let $x = 2s, s \in R$, thus $y = 3s$	Let $x = s, s \in R$, thus $y = s$
Eigenvectors are $s\begin{bmatrix} 2 \\ 3 \end{bmatrix} \rightarrow$ an eigenvector is $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$.	Eigenvectors are $s\begin{bmatrix} 1 \\ 1 \end{bmatrix} \rightarrow$ an eigenvector is $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

The matrix has eigenvalues $1, 2$ with corresponding eigenvectors $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Example 3: W2002 P1 Q11 EITHER

The vector **e** is an eigenvector of the square matrix **G**. Show that

(i) **e** is an eigenvector of $\mathbf{G} + k\mathbf{I}$, where k is a scalar and **I** is an identity matrix,

(ii) **e** is an eigenvector of $\mathbf{G}^2$.

[5]

Find the eigenvalues, and corresponding eigenvectors, of the matrices **A** and $\mathbf{B}^2$, where

$$\mathbf{A} = \begin{pmatrix} 3 & -3 & 0 \\ 1 & 0 & 1 \\ -1 & 3 & 2 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} -5 & -3 & 0 \\ 1 & -8 & 1 \\ -1 & 3 & -6 \end{pmatrix}. \qquad [9]$$

(i)

$$\mathbf{G}\mathbf{e} = \lambda\mathbf{e}$$

$$\begin{aligned}(\mathbf{G} + k\mathbf{I})\mathbf{e} &= \mathbf{G}\mathbf{e} + k\mathbf{I}\mathbf{e}\\ &= \lambda\mathbf{e} + k\mathbf{e} = (\lambda + k)\mathbf{e}\end{aligned}$$

Thus matrix $(\mathbf{G} + k\mathbf{I})$ has eigenvector **e** with corresponding eigenvalue $\lambda + k$.

(ii)

$$\begin{aligned}\mathbf{G}^2\mathbf{e} &= \mathbf{G}(\mathbf{G}\mathbf{e})\\ &= \mathbf{G}(\lambda\mathbf{e})\\ &= \lambda(\mathbf{G}\mathbf{e})\\ &= \lambda(\lambda\mathbf{e}) = \lambda^2\mathbf{e}\end{aligned}$$

Thus matrix $\mathbf{G}^2$has eigenvector **e** with corresponding eigenvalue λ^2.

Strategy: We find eigenvalues and eigenvectors of matrix B then use the above two rules to give the final answer.

Eigenvalues:

$$\begin{vmatrix}-5-\lambda & -3 & 0\\ 1 & -8-\lambda & 1\\ -1 & 3 & -6-\lambda\end{vmatrix}=0$$

$$(-5-\lambda)\begin{vmatrix}-8-\lambda & 1\\ 3 & -6-\lambda\end{vmatrix}+3\begin{vmatrix}1 & 1\\ -1 & -6-\lambda\end{vmatrix}=0$$

$$(5+\lambda)[(8+\lambda)(6+\lambda)-3]-3(-6-\lambda+1)=0$$

$$(5+\lambda)(\lambda^2+14\lambda+45)-3(-5-\lambda)=0$$

$$(5+\lambda)(\lambda^2+14\lambda+45)-3(-5-\lambda)=0$$

$$(5+\lambda)(\lambda^2+14\lambda+48)=0$$

$$(5+\lambda)(\lambda+6)(\lambda+8)=0,\quad \lambda=-5,-6,-8$$

When $\lambda=-5$:

$$\begin{bmatrix}-5 & -3 & 0\\ 1 & -8 & 1\\ -1 & 3 & -6\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=-5\begin{bmatrix}x\\ y\\ z\end{bmatrix}$$

$-3y=-5y$ $\quad x-3y+z=0$ $\quad -x+3y-z=0$

$y=0$ $\quad x=-z$ $\quad x=-z$

Let $x=s, s\in R$, thus $z=-s$.

Eigenvectors are $s\begin{bmatrix}1\\0\\-1\end{bmatrix}\rightarrow$ an eigenvector is $\begin{bmatrix}1\\0\\-1\end{bmatrix}$.

When $\lambda=-6$:

$$\begin{bmatrix}-5 & -3 & 0\\ 1 & -8 & 1\\ -1 & 3 & -6\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=-6\begin{bmatrix}x\\ y\\ z\end{bmatrix}$$

$x-3y=0$ $\quad x-2y+z=0$ $\quad -x+3y=0$

$x=3y$

Let $y=s, s\in R$, thus $x=3s.\ z=2s-3s=-s$

Eigenvectors are $s\begin{bmatrix}3\\1\\-1\end{bmatrix}\rightarrow$ an eigenvector is $\begin{bmatrix}3\\1\\-1\end{bmatrix}$.

When $\lambda=-8$:

$$\begin{bmatrix}-5 & -3 & 0\\ 1 & -8 & 1\\ -1 & 3 & -6\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=-8\begin{bmatrix}x\\ y\\ z\end{bmatrix}$$

$3x-3y=0$ $\quad x+z=0$ $\quad -x+3y+2z=0$

Let $x=s, s\in R$, thus $y=s, z=-s.\ -s+3s-2s=0$

Eigenvectors are $s\begin{bmatrix}1\\1\\-1\end{bmatrix}\rightarrow$ an eigenvector is $\begin{bmatrix}1\\1\\-1\end{bmatrix}$.

B has eigenvalues $-5,-6,-8$ with corresponding eigenvectors $\begin{bmatrix}1\\0\\-1\end{bmatrix},\begin{bmatrix}3\\1\\-1\end{bmatrix},\begin{bmatrix}1\\1\\-1\end{bmatrix}$.

Since $\mathbf{A}=\mathbf{B}+8\mathbf{I}$, using (i), **A** has eigenvalues $3, 2, 0$ with corresponding eigenvectors

$\begin{bmatrix}1\\0\\-1\end{bmatrix}, \begin{bmatrix}3\\1\\-1\end{bmatrix}, \begin{bmatrix}1\\1\\-1\end{bmatrix}$.

Using (ii), $\mathbf{B}^2$ has eigenvalues $25, 36, 64$ with corresponding eigenvectors $\begin{bmatrix}1\\0\\-1\end{bmatrix}, \begin{bmatrix}3\\1\\-1\end{bmatrix}, \begin{bmatrix}1\\1\\-1\end{bmatrix}$.

Example 4: W2009 P1 Q11 OR

One of the eigenvalues of the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & -4 & 6 \\ 2 & -4 & 2 \\ -3 & 4 & a \end{pmatrix}$$

is -2. Find the value of a. [3]

Another eigenvalue of $\mathbf{A}$ is -5. Find eigenvectors $\mathbf{e}_1$ and $\mathbf{e}_2$ corresponding to the eigenvalues -2 and -5 respectively. [3]

The linear space spanned by $\mathbf{e}_1$ and $\mathbf{e}_2$ is denoted by V.

(i) Prove that, for any vector $\mathbf{x}$ belonging to V, the vector $\mathbf{Ax}$ also belongs to V. [4]

(ii) Find a non-zero vector which is perpendicular to every vector in V, and determine whether it is an eigenvector of $\mathbf{A}$. [4]

When $\lambda = -2$:

$$\begin{vmatrix} 1+2 & -4 & 6 \\ 2 & -4+2 & 2 \\ -3 & 4 & a+2 \end{vmatrix} = 0$$

$$-3\begin{vmatrix} -4 & 6 \\ -2 & 2 \end{vmatrix} - 4\begin{vmatrix} 3 & 6 \\ 2 & 2 \end{vmatrix} + (a+2)\begin{vmatrix} 3 & -4 \\ 2 & -2 \end{vmatrix} = 0$$

$$-3(4) - 4(-6) + (a+2)(2) = 0$$

$$2a + 16 = 0, a = -8$$

$$\begin{bmatrix} 1 & -4 & 6 \\ 2 & -4 & 2 \\ -3 & 4 & -8 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = -2\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$3x - 4y + 6z = 0$ $\quad 2x - 2y + 2z = 0$ $\quad -3x + 4y - 6z = 0$

$4y = 3x + 6z$ $\quad x + z = y$

$4x + 4z = 3x + 6z$

$x = 2z$

Let $z = s, s \in R$, thus $x = 2s, y = 2s + s = 3s$

Eigenvectors are $s\begin{bmatrix}2\\3\\1\end{bmatrix}$ → an eigenvector is $\begin{bmatrix}2\\3\\1\end{bmatrix}$.

When $\lambda = -5$:

$$\begin{bmatrix} 1 & -4 & 6 \\ 2 & -4 & 2 \\ -3 & 4 & -8 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = -5\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$6x - 4y + 6z = 0$ $\quad 2x + y + 2z = 0$ $\quad -3x + 4y - 3z = 0$

$y = 0, x + z = 0$, let $z = -s, s \in R$, thus $x = s$.

Eigenvectors are $s\begin{bmatrix}1\\0\\-1\end{bmatrix}$ → an eigenvector is $\begin{bmatrix}1\\0\\-1\end{bmatrix}$.

(i) Vectors in *V*, **x** have the form $p\mathbf{e_1} + q\mathbf{e_2}$ for real numbers *p* and *q*.

$$\mathbf{Ax} = A(p\mathbf{e_1} + q\mathbf{e_2}) = p\mathbf{Ae_1} + q\mathbf{Ae_2} = -2p\mathbf{e_1} - 5q\mathbf{e_2}$$

Since $-2p$ and $-5q$ are real numbers, vector **Ax** is also in *V*.

(ii)
$$\begin{bmatrix}2\\3\\1\end{bmatrix} \times \begin{bmatrix}1\\0\\-1\end{bmatrix} = \begin{bmatrix}0-3\\1+2\\0-3\end{bmatrix} = -3\begin{bmatrix}1\\-1\\1\end{bmatrix}.$$

$\mathbf{e_3}$: $\boldsymbol{i} - \boldsymbol{j} + \boldsymbol{k}$ is perpendicular to every vector in *V*.

$$\begin{bmatrix}1 & -4 & 6\\2 & -4 & 2\\-3 & 4 & -8\end{bmatrix}\begin{bmatrix}1\\-1\\1\end{bmatrix} = \begin{bmatrix}1+4+6\\2+4+2\\-3-4-8\end{bmatrix} = \begin{bmatrix}11\\8\\-15\end{bmatrix}$$

Since $\mathbf{Ae_3}$ is not a scalar multiple of $\mathbf{e_3}$, $\mathbf{e_3}$ is not an eigenvector of **A**.

B) Diagonalisation

- LO: express a matrix in the form $\mathbf{QDQ}^{-1}$, where **D** is a diagonal matrix of eigenvalues and **Q** is a matrix whose columns are eigenvectors, and use this expression, e.g. in calculating powers of matrices.

  - A $n \times n$ matrix **A** is said to be diagonalizable if there exists a non-singular matrix **P** and a diagonal matrix **D** such that $\mathbf{P^{-1}AP} = \mathbf{D}$. To do so, **A** must have *n* linearly independent eigenvectors.
  - This can be rearranged into $\mathbf{A} = \mathbf{PDP^{-1}}$ by pre-multiplying both sides with **P** and post-multiplying with $\mathbf{P^{-1}}$.

1. This can be used to find powers of matrix A, as if $\mathbf{A} = \mathbf{PDP^{-1}}$ then $\mathbf{A}^n = \mathbf{PD}^n\mathbf{P^{-1}}$.

2. Also for a diagonal matrix $\mathbf{D} = \begin{bmatrix}\lambda_1 & 0 & 0 & 0\\0 & \lambda_2 & 0 & 0\\... & ... & ... & 0\\0 & 0 & 0 & \lambda_n\end{bmatrix}$, $\mathbf{D}^m = \begin{bmatrix}\lambda_1^m & 0 & 0 & 0\\0 & \lambda_2^m & 0 & 0\\... & ... & ... & 0\\0 & 0 & 0 & \lambda_n^m\end{bmatrix}$

(Both statements can be proven using mathematical induction and is left as an exercise)

  - If **A** has eigenvalues $\lambda_1,\ \lambda_2, \dots, \lambda_n$ with corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$, **P** is a matrix formed using its eigenvectors as columns and **D** is a matrix formed using its eigenvalues as elements, as below:

$$\text{If} = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad ... \quad \mathbf{v}_n],\ \mathbf{D} = \begin{bmatrix}\lambda_1 & 0 & 0 & 0\\0 & \lambda_2 & 0 & 0\\... & ... & ... & 0\\0 & 0 & 0 & \lambda_n\end{bmatrix}.$$

  - The arrangement of the columns in the matrix can vary, as long as the eigenvalue and its corresponding eigenvector are in the same column (if λ_2 is in column 4 so is $\mathbf{v}_2$). This method is shown in below (observe how both methods give the same results).

Example 1: Matrix **A** has eigenvalues 1, 2 with corresponding eigenvectors $\begin{bmatrix}2\\3\end{bmatrix}$, $\begin{bmatrix}1\\1\end{bmatrix}$. Find

(i) **A** (ii) $\mathbf{A} + \mathbf{A}^2$ (iii) $\mathbf{A}^n$, where *n* is a positive integer.

Method 1: Let $\mathbf{P} = \begin{bmatrix}2 & 1\\3 & 1\end{bmatrix}$, $\mathbf{D} = \begin{bmatrix}1 & 0\\0 & 2\end{bmatrix}$.

Method 2: Let $\mathbf{P} = \begin{bmatrix}1 & 2\\1 & 3\end{bmatrix}$, $\mathbf{D} = \begin{bmatrix}2 & 0\\0 & 1\end{bmatrix}$.

(i) $\mathbf{P}^{-1} = \frac{1}{2-3}\begin{bmatrix} 1 & -1 \\ -3 & 2 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 3 & -2 \end{bmatrix}$

$\mathbf{A} = \mathbf{PDP}^{-1}$

$= \begin{bmatrix} 2 & 1 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 3 & -2 \end{bmatrix}$

$= \begin{bmatrix} 2 & 2 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 3 & -2 \end{bmatrix}$

$= \begin{bmatrix} -2+6 & 2-4 \\ -3+6 & 3-4 \end{bmatrix} = \begin{bmatrix} 4 & -2 \\ 3 & -1 \end{bmatrix}$

(ii) $\mathbf{A}^2 = \mathbf{PD}^2\mathbf{P}^{-1}$

$= \begin{bmatrix} 2 & 1 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 1^2 & 0 \\ 0 & 2^2 \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 3 & -2 \end{bmatrix}$

$= \begin{bmatrix} 2 & 4 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 3 & -2 \end{bmatrix}$

$= \begin{bmatrix} -2+12 & 2-8 \\ -3+12 & 3-8 \end{bmatrix} = \begin{bmatrix} 10 & -6 \\ 9 & -5 \end{bmatrix}$

$\mathbf{A} + \mathbf{A}^2 = \begin{bmatrix} 4 & -2 \\ 3 & -1 \end{bmatrix} + \begin{bmatrix} 10 & -6 \\ 9 & -5 \end{bmatrix} = \begin{bmatrix} 14 & -8 \\ 12 & -6 \end{bmatrix}$

(iii) $\mathbf{A}^n = \mathbf{PD}^n\mathbf{P}^{-1}$

$= \begin{bmatrix} 2 & 1 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 1^n & 0 \\ 0 & 2^n \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 3 & -2 \end{bmatrix}$

$= \begin{bmatrix} 2 & 2^n \\ 3 & 2^n \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 3 & -2 \end{bmatrix}$

$= \begin{bmatrix} -2+3.2^n & 2-2.2^n \\ -3+3.2^n & 3-2.2^n \end{bmatrix}$

(i) $\mathbf{P}^{-1} = \frac{1}{3-2}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$

$\mathbf{A} = \mathbf{PDP}^{-1}$

$= \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$

$= \begin{bmatrix} 2 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$

$= \begin{bmatrix} 6-2 & -4+2 \\ 6-3 & -4+3 \end{bmatrix} = \begin{bmatrix} 4 & -2 \\ 3 & -1 \end{bmatrix}$

(ii) $\mathbf{A}^2 = \mathbf{PD}^2\mathbf{P}^{-1}$

$= \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 2^2 & 0 \\ 0 & 1^2 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$

$= \begin{bmatrix} 4 & 2 \\ 4 & 3 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$

$= \begin{bmatrix} 12-2 & -8+2 \\ 12-3 & -8+3 \end{bmatrix} = \begin{bmatrix} 10 & -6 \\ 9 & -5 \end{bmatrix}$

(iii) $\mathbf{A}^n = \mathbf{PD}^n\mathbf{P}^{-1}$

$= \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 2^n & 0 \\ 0 & 1^n \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$

$= \begin{bmatrix} 2^n & 2 \\ 2^n & 3 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$

$= \begin{bmatrix} 3.2^n-2 & -2.2^n+2 \\ 3.2^n-3 & -2.2^n+3 \end{bmatrix}$

Example 2: W2004 P1 Q12 OR

The matrix $\mathbf{A}$ has λ as an eigenvalue with $\mathbf{e}$ as a corresponding eigenvector. Show that if $\mathbf{A}$ is non-singular then

(i) $\lambda \neq 0$, [2]

(ii) the matrix $\mathbf{A}^{-1}$ has λ^{-1} as an eigenvalue with $\mathbf{e}$ as a corresponding eigenvector. [2]

The matrices $\mathbf{A}$ and $\mathbf{B}$ are given by

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 2 \\ 0 & -2 & 4 \\ 0 & 0 & -3 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = (\mathbf{A} + 4\mathbf{I})^{-1}.$$

Find a non-singular matrix $\mathbf{P}$, and a diagonal matrix $\mathbf{D}$, such that $\mathbf{B} = \mathbf{PDP}^{-1}$. [10]

(i) If $\lambda = 0$: $\det(\mathbf{A} - 0\mathbf{I}) = 0 \rightarrow \det(\mathbf{A}) = 0$, thus $\mathbf{A}$ is singular.
Contrapositive of the statement: If A is non-singular, then $\lambda \neq 0$.
Note: Statements (if p then q) and (if not q then not p) are identical.

(ii) $\mathbf{Ae} = \lambda\mathbf{e} \rightarrow \mathbf{A}^{-1}\mathbf{Ae} = \mathbf{A}^{-1}\lambda\mathbf{e}$

$\mathbf{Ie} = \lambda\mathbf{A}^{-1}\mathbf{e}$

$\mathbf{A}^{-1}\mathbf{e} = \lambda^{-1}\mathbf{Ie} = \lambda^{-1}\mathbf{e}$

Strategy: Find eigenvalues and eigenvectors of $\mathbf{A} + 4\mathbf{I}$, then use (ii). [or find eigenvalues and eigenvectors for $\mathbf{A}$ then increase the eigenvalues by 4.]

Eigenvalues:

$$\begin{vmatrix} 5-\lambda & -1 & 2 \\ 0 & 2-\lambda & 4 \\ 0 & 0 & 1-\lambda \end{vmatrix} = 0$$

$$(1-\lambda)\begin{vmatrix} 5-\lambda & -1 \\ 0 & 2-\lambda \end{vmatrix} = 0$$

$$(1-\lambda)(5-\lambda)(2-\lambda) = 0, \quad \lambda = 1, 2, 5$$

When $\lambda = 1$: $\begin{bmatrix} 5 & -1 & 2 \\ 0 & 2 & 4 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

$4x - y + 2z = 0 \qquad y + 4z = 0 \qquad z = z$

Let $z = -2s, s \in R$, thus $y = 8s, x = \frac{1}{4}(8s + 4s) = 3s$.

Eigenvectors are $s\begin{bmatrix} 3 \\ 8 \\ -2 \end{bmatrix}$ → an eigenvector is $\begin{bmatrix} 3 \\ 8 \\ -2 \end{bmatrix}$.

When $\lambda = 2$: $\begin{bmatrix} 5 & -1 & 2 \\ 0 & 2 & 4 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = 2\begin{bmatrix} x \\ y \\ z \end{bmatrix}$

$3x - y + 2z = 0 \qquad 2y + 4z = 2y \qquad z = 2z, z = 0$

$y = 3x$

Let $y = 3s, s \in R,$ thus $x = s$.

Eigenvectors are $s\begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}$ → an eigenvector is $\begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}$.

When $\lambda = 5$: $\begin{bmatrix} 5 & -1 & 2 \\ 0 & 2 & 4 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = 5\begin{bmatrix} x \\ y \\ z \end{bmatrix}$

$5x - y + 2z = 5x \qquad 3y + 4z = 0 \qquad z = 5z, z = 0$

Let $x = s, s \in R \qquad y = 0$

Eigenvectors are $s\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ → an eigenvector is $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

Using (ii), matrix **B** has eigenvalues, $1^{-1},\ 2^{-1},\ 5^{-1}$ with corresponding eigenvectors $\begin{bmatrix} 3 \\ 8 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. Thus $\mathbf{P} = \begin{bmatrix} 3 & 1 & 1 \\ 8 & 3 & 0 \\ -2 & 0 & 0 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.2 \end{bmatrix}$

Example 3: W2005 P1 Q10

It is given that the eigenvalues of the matrix **M**, where

$$\mathbf{M} = \begin{pmatrix} 4 & 1 & -1 \\ -4 & -1 & 4 \\ 0 & -1 & 5 \end{pmatrix},$$

are 1, 3, 4. Find a set of corresponding eigenvectors. [4]

Write down a matrix **P** and a diagonal matrix **D** such that

$$\mathbf{M}^n = \mathbf{PDP}^{-1},$$

where n is a positive integer. [2]

Find $\mathbf{P}^{-1}$ and deduce that

$$\lim_{n\to\infty} 4^{-n}\mathbf{M}^n = \begin{pmatrix} -\frac{1}{3} & 0 & -\frac{1}{3} \\ \frac{4}{3} & 0 & \frac{4}{3} \\ \frac{4}{3} & 0 & \frac{4}{3} \end{pmatrix}. \qquad [5]$$

When $\lambda = 1$: $\begin{bmatrix} 4 & 1 & -1 \\ -4 & -1 & 4 \\ 0 & -1 & 5 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

$3x + y - z = 0 \qquad -4x - 2y + 4z = 0 \qquad -y + 4z = 0$

Let $z = -s, s \in R$, thus $y = -4s, x = \frac{1}{3}(-s + 4s) = s$

Eigenvectors are $s\begin{bmatrix}1\\-4\\-1\end{bmatrix} \rightarrow$ an eigenvector is $\begin{bmatrix}1\\-4\\-1\end{bmatrix}$.

When $\lambda = 3$:

$$\begin{bmatrix}4 & 1 & -1\\-4 & -1 & 4\\0 & -1 & 5\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = 3\begin{bmatrix}x\\y\\z\end{bmatrix}$$

$x + y - z = 0 \qquad -4x - 4y + 4z = 0 \qquad -y + 2z = 0$

Let $z = -s, s \in R$, thus $y = -2s, x = -s + 2s = s$

Eigenvectors are $s\begin{bmatrix}1\\-2\\-1\end{bmatrix} \rightarrow$ an eigenvector is $\begin{bmatrix}1\\-2\\-1\end{bmatrix}$.

When $\lambda = 4$:

$$\begin{bmatrix}4 & 1 & -1\\-4 & -1 & 4\\0 & -1 & 5\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = 4\begin{bmatrix}x\\y\\z\end{bmatrix}$$

$y - z = 0 \qquad -4x - 5y + 4z = 0 \qquad -y + z = 0$

Let $z = -4s, s \in R$, thus $y = -4s, x = \frac{1}{4}(20s - 16s) = s$

Eigenvectors are $s\begin{bmatrix}1\\-4\\-4\end{bmatrix} \rightarrow$ an eigenvector is $\begin{bmatrix}1\\-4\\-4\end{bmatrix}$.

$\mathbf{M}^n$ has eigenvalues 1, 3^n, 4^n and corresponding eigenvectors $\begin{bmatrix}1\\-4\\-1\end{bmatrix}, \begin{bmatrix}1\\-2\\-1\end{bmatrix}, \begin{bmatrix}1\\-4\\-4\end{bmatrix}$.

Thus $\mathbf{P} = \begin{bmatrix}1 & 1 & 1\\-4 & -2 & -4\\-1 & -1 & -4\end{bmatrix}, \qquad \mathbf{D} = \begin{bmatrix}1 & 0 & 0\\0 & 3^n & 0\\0 & 0 & 4^n\end{bmatrix}$

$$\det(\mathbf{P}) = \begin{vmatrix}-2 & -4\\-1 & -4\end{vmatrix} - \begin{vmatrix}-4 & -4\\-1 & -4\end{vmatrix} + \begin{vmatrix}-4 & -2\\-1 & -1\end{vmatrix} = 4 - 12 + 2 = -6$$

Matrix of Cofactors: $\begin{bmatrix}\begin{vmatrix}-2 & -4\\-1 & -4\end{vmatrix} & -\begin{vmatrix}-4 & -4\\-1 & -4\end{vmatrix} & \begin{vmatrix}-4 & -2\\-1 & -1\end{vmatrix}\\ -\begin{vmatrix}1 & 1\\-1 & -4\end{vmatrix} & \begin{vmatrix}1 & 1\\-1 & -4\end{vmatrix} & -\begin{vmatrix}1 & 1\\-1 & -1\end{vmatrix}\\ \begin{vmatrix}1 & 1\\-2 & -4\end{vmatrix} & -\begin{vmatrix}1 & 1\\-4 & -4\end{vmatrix} & \begin{vmatrix}1 & 1\\-4 & -2\end{vmatrix}\end{bmatrix} = \begin{bmatrix}4 & -12 & 2\\3 & -3 & 0\\-2 & 0 & 2\end{bmatrix}$

$$\mathbf{P}^{-1} = \frac{1}{-6}\begin{bmatrix}4 & -12 & 2\\3 & -3 & 0\\-2 & 0 & 2\end{bmatrix}^T = \frac{1}{6}\begin{bmatrix}-4 & -3 & 2\\12 & 3 & 0\\-2 & 0 & -2\end{bmatrix}$$

$$4^{-n}\mathbf{M}^n = \frac{1}{6}\begin{bmatrix}1 & 1 & 1\\-4 & -2 & -4\\-1 & -1 & -4\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\0 & 3^n & 0\\0 & 0 & 4^n\end{bmatrix}\begin{bmatrix}-4 & -3 & 2\\12 & 3 & 0\\-2 & 0 & -2\end{bmatrix}(4^{-n})$$

$$= \frac{1}{6}\begin{bmatrix}1 & 1 & 1\\-4 & -2 & -4\\-1 & -1 & -4\end{bmatrix}\begin{bmatrix}-4 & -3 & 2\\12.3^n & 3.3^n & 0\\-2.4^n & 0 & -2.4^n\end{bmatrix}(4^{-n})$$

$$= \frac{1}{6}\begin{bmatrix}1 & 1 & 1\\-4 & -2 & -4\\-1 & -1 & -4\end{bmatrix}\begin{bmatrix}-4.4^{-n} & -3.4^{-n} & 2.4^{-n}\\12(0.75^n) & 3(0.75^n) & 0\\-2 & 0 & -2\end{bmatrix}$$

As $n \rightarrow \infty$, $4^{-n} \rightarrow 0$, $0.75^n \rightarrow 0$,

$$\lim_{n\to\infty} 4^{-n}\mathbf{M}^n = \frac{1}{6}\begin{bmatrix}1 & 1 & 1\\-4 & -2 & -4\\-1 & -1 & -4\end{bmatrix} = \frac{1}{6}\begin{bmatrix}-2 & 0 & -2\\8 & 0 & 8\\8 & 0 & 8\end{bmatrix} = \begin{bmatrix}-\frac{1}{3} & 0 & -\frac{1}{3}\\ \frac{4}{3} & 0 & \frac{4}{3}\\ \frac{4}{3} & 0 & \frac{4}{3}\end{bmatrix}$$

FURTHER MATHEMATICS PAPER 2 MECHANICS

CHAPTER 10: MOMENTUM AND IMPULSE

10.1 Pre-requisite Knowledge

A) Paper 4: Kinematics and Energy

- Equations of constant acceleration:
  $v = u + at,\ \ s = ut + \frac{1}{2}at^2,\ \ s = \frac{1}{2}(u+v)t,\ \ v^2 = u^2 + 2as,\ \ s = vt - \frac{1}{2}at^2$
- Kinetic energy, $E_k = \frac{1}{2}mv^2$, possessed by a moving object.
- Gravitational potential energy, $E_p = mgh$, possessed by any object that does work against gravity.
- Work-energy principle: The work done by the force acting on an object, less the work done against resistance, is equal to the gain in kinetic and potential energy of the object.
- Conservation of energy: For an object moving along a path, if there is no work done by external forces other than the force of gravity, the sum of potential energy and kinetic energy is constant.

B) Paper 5: Motion of a projectile

If a projectile is launched with initial velocity v with angle θ from the horizontal:

- It moves horizontally with constant velocity (assuming zero air resistance) and moves vertically with constant acceleration (free-fall).
- Displacement equations: $s_x = vt\cos\theta\,, s_y = vt\sin\theta - \frac{1}{2}gt^{2`}$
- Greatest height, $h = \frac{u^2 sin^2\theta}{2g}$.
- Range on horizontal ground, $= \frac{u^2\sin 2\theta}{g}$
- $r_{max} = \frac{u^2}{g}$, when $\theta = 45^o$.
- Time to return to original height, $= \frac{2u\sin\theta}{g}$.
- Equation of trajectory, $y = x\tan\theta - \frac{gx^2}{2v^2cos^2\theta}$.

Example 1: W2003 Mathematics P5 Q5

A stone is projected from a point on horizontal ground with a speed of 20 m s^{-1} at an angle of $\alpha°$ above the horizontal. The stone is moving horizontally when it hits a vertical wall at a point 7.2 m above the ground.

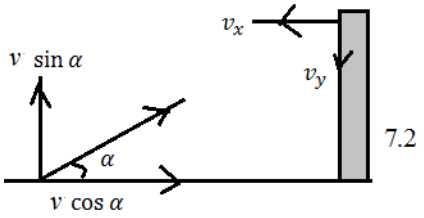
(i) Find the value of α. [4]

After rebounding at right angles from the wall the speed of the stone is halved. Find

(ii) the distance from the wall of the point at which the stone hits the ground, [4]

(iii) the angle which the direction of motion of the stone makes with the horizontal, immediately before the stone hits the ground. [3]

(i)

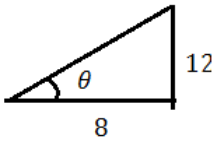


$$v_y = 20\sin\alpha$$
$$v^2 = u^2 - 2gs$$
$$400\sin^2\alpha = 20(7.2)$$
$$\sin\alpha = \sqrt{\frac{144}{400}} \rightarrow \alpha = 36.9^o$$

(ii)

$$s = ut + \frac{1}{2}at^2$$
$$7.2 = 5t^2 \Rightarrow t = 1.2s$$
$$v_x = 20\cos\alpha\left(\frac{1}{2}\right) = 8\ ms^{-1}$$
$$s_x = v_x t = 9.6\ m$$

(iii)



$$v_y = 10\ (1.2) = 12\ ms^{-1}$$
$$\theta = \tan^{-1}\frac{12}{8} = 56.3^o$$

Example 2: S2002 Mathematics P5 Q7

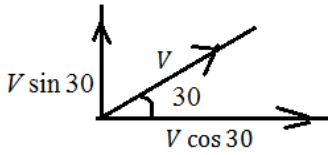
A ball is projected from a point O with speed V m s^{-1}, at an angle of 30° above the horizontal. At time T s after projection, the ball passes through the point A, whose horizontal and vertically upward displacements from O are 10 m and 2 m respectively.

(i) By using the equation of the trajectory, or otherwise, find the value of V. [3]

(ii) Find the value of T. [2]

(iii) Find the angle that the direction of motion of the ball at A makes with the horizontal. [4]

(i)

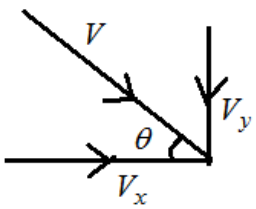


$$2 = 10\tan 30 - \frac{100g}{2V^2 cos^2 30}$$
$$V^2 = \frac{50g}{cos^2 30(10\tan 30 - 2)}$$
$$V = 13.3\ ms^{-1}$$

(ii)

$$V\cos 30\,.\,T = 10$$
$$T = \frac{10}{13.3\cos 30} = 0.869\ s$$

(iii)



$$V_y = V\sin 30 - gT$$
$$= 13.3\sin 30 - 8.69 = -2.04$$
$$V_x = 13.3\cos 30 = 11.52$$
$$\tan\theta = \left|\frac{V_y}{V_x}\right| \rightarrow \theta = 10.0^o$$

10.2 Momentum, Impulse and Collisions

- LO: recall and use the definition of linear momentum, and show understanding of its vector nature (in one dimension only);
- LO: recall Newton's experimental law and the coefficient of restitution, the property $0 \leq e \leq 1$, and the meaning of the terms 'perfectly elastic' ($e = 1$) and 'inelastic' ($e = 0$);
- LO: use conservation of linear momentum and/or Newton's experimental law to solve problems that may be modelled as the direct impact of two smooth spheres or the impact of a smooth sphere with a fixed surface;
- LO: recall and use the definition of the impulse of a constant force, and relate the impulse acting on a particle to the change of momentum of the particle (in one dimension only).

- For an object of mass m moving with velocity v, momentum = mv.
- Impulse received by an object is equal to the increase in momentum of the object in the direction of the impulse.
- The impulse-momentum equation: $I = Ft = mv - mu$
- Conservation of momentum: If two objects moving along the same line collide with each other, the total momentum of the two objects is the same after the collision as it was before.
- By Newton's 3rd law, if the first body receive an impulse of I, the second body receives an impulse of $-I$.
- Newton's law of impact states that if two objects moving in the same straight line collide, the ratio $\frac{\text{seperation speed}}{\text{approach speed}}$ is constant. This constant is known as coefficient of restitution and it depends on the shape of the objects and the materials they consist of.
- The coefficient of restitution, e is always between 0 and 1.
  If $e = 1$, the collision is perfectly elastic. The kinetic energy before the collision and after the collision will be constant.

  If $e = 0$, the collision is inelastic, and the particles stick together / coalesce.
- Algebraic form of Newton's law:
$$v_2 - v_1 = -e(u_2 - u_1) \rightarrow e = \frac{v_2 - v_1}{u_1 - u_2}$$
- Newton's law still applies when a particle hits and rebounds from a fixed flat surface.

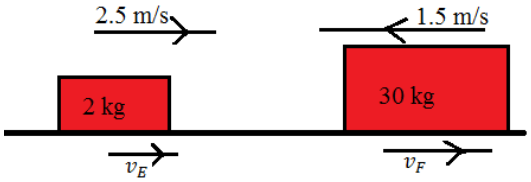
Problem solving strategies:

1. Draw a free body diagram
2. Analyse every collision by using the conservation of momentum equation
3. Analyse every collision by using the equation for coefficient of restitution
4. Solve simultaneous equations to find unknown velocities
5. Use kinetic energy / impulse / projectile equations if necessary

Example 1: W2003 P2 Q2

An empty shopping trolley E of mass 2 kg is moving at a speed of 2.5 m s^{-1} on a smooth horizontal floor towards a full trolley F of mass 30 kg. Just before they collide, F is moving at 1.5 m s^{-1} towards E. The collision is modelled as a direct impact between two particles with coefficient of restitution 0.2. Find the speed and the direction of motion of E immediately after the collision. [5]

During the impact the trolleys are in contact for 0.1 s, and the force between them is constant. Find the magnitude of this force. [2]



Conservation of momentum:

$$2(2.5) - 1.5(30) = 2v_E + 30v_F$$
$$2v_E + 30v_F = -40 - (1)$$

Restitution:

$$\frac{v_F - v_E}{2.5 + 1.5} = 0.2$$
$$v_F - v_E = 0.8$$
$$30v_F - 30v_E = 24 - (2)$$

$$(1) - (2): 32v_E = -16, v_E = -2$$

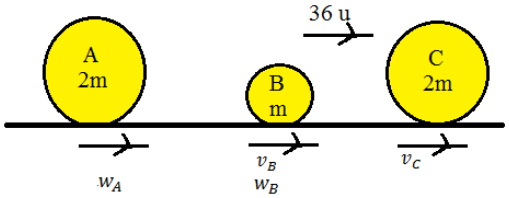
E moves with speed $2ms^{-1}$ in the opposite direction which it is initially moving.

$$Ft = \Delta p$$
$$0.1F = 2(2.5 + 2)$$
$$F = 90N$$

Example 2: S2005 P2 Q3

Three balls, A, B and C, of the same size but of masses $2m$, m and $2m$ respectively, lie at rest in a straight line on a smooth horizontal table. The ball B, which is between A and C, is projected towards C with speed $36u$. The coefficients of restitution between A and B, and between B and C, are each $\frac{3}{4}$. Show that the second collision is the last. [7]

Express the total loss of kinetic energy of the system as a percentage of the original kinetic energy. [2]



Collision 1: Between B and C

Conservation of momentum:

$$36mu = mv_B + 2mv_C$$
$$v_B + 2v_C = 36u - (1)$$

Restitution:

$$\frac{v_C - v_B}{36u - 0} = \frac{3}{4}$$
$$v_C - v_B = 27u - (2)$$

$$(1) + (2): 3v_C = 63u, v_c = 21u$$
$$v_B = 36u - 42u = -6u$$

Collision 2: Between A and B

Conservation of momentum:

$$-6mu = mw_B + 2mw_A$$
$$w_B + 2w_A = -6u - (3)$$

Restitution:

$$\frac{w_B - w_A}{0 + 6u} = \frac{3}{4}$$

$$w_B - w_A = \frac{9}{2}u - (4)$$

$$(3) + 2(4): 3w_B = 3u, w_B = u$$

$$w_A = u - 4.5u = -3.5u$$

After the second collision, A moves to the left while B and C moves to the right.

But since $v_C > v_B$, no more collisions will occur between the three particles.

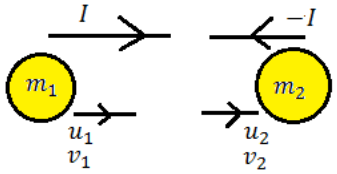
$$\text{Initial KE} = \frac{1}{2}m(36u)^2 = 648mu^2$$

$$\text{Final KE} = \frac{1}{2}(2m)\left(\frac{7}{2}u\right)^2 + \frac{1}{2}mu^2 + \frac{1}{2}(2m)(21u)^2 = 453.75mu^2$$

$$\text{Loss in KE} = \frac{648 - 453.75}{648} \times 100\% = 30.0\% \ (3\ s.f.)$$

Example 3: S2007 P2 Q1

Particles of masses m_1 and m_2 lie at rest on a smooth horizontal plane. Each particle is given a horizontal impulse of magnitude I towards the other particle so that the particles collide directly. Given that the collision is perfectly elastic, find the speeds of the particles after the collision. [6]



$$I = m_1u_1 = -m_2u_2$$

Conservation of momentum:

$$m_1v_1 + m_2v_2 = 0$$

$$v_1 = -\frac{m_2v_2}{m_1} - (1)$$

Restitution:

$$e = 1: v_2 - v_1 = u_1 - u_2$$

$$v_2 - v_1 = \frac{I}{m_1} + \frac{I}{m_2} - (2)$$

$$(1) into\ (2): v_2 - \frac{m_2v_2}{m_1} = \frac{I}{m_1} + \frac{I}{m_2}$$

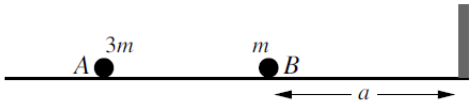
$$v_2\left(1 + \frac{m_2}{m_1}\right) = I\left(\frac{m_1 + m_2}{m_1m_2}\right)$$

$$v_2 = I\left(\frac{m_1 + m_2}{m_1m_2}\right)\left(\frac{m_1}{m_1 + m_2}\right) = \frac{I}{m_2}$$

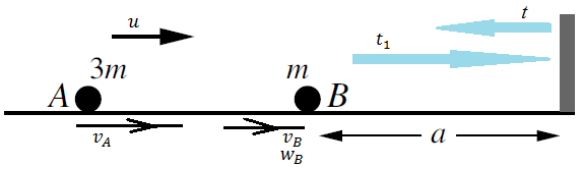
$$v_1 = -\frac{m_2}{m_1}\left(\frac{I}{m_2}\right) = -\frac{I}{m_1}$$

After the collision, particle 1 has speed $\frac{I}{m_1}$ and particle 2 has speed $\frac{I}{m_2}$.

Example 4: S2010 P21 Q3



Two perfectly elastic small smooth spheres A and B have masses $3m$ and m respectively. They lie at rest on a smooth horizontal plane with B at a distance a from a smooth vertical barrier. The line of centres of the spheres is perpendicular to the barrier, and B is between A and the barrier (see diagram). Sphere A is projected towards sphere B with speed u and, after the collision between the spheres, B hits the barrier. The coefficient of restitution between B and the barrier is $\frac{1}{2}$. Find the speeds of A and B immediately after they first collide, and the distance from the barrier of the point where they collide for the second time. [9]



Collision 1: Between A and B

Conservation of momentum:

$$3mu = 3mv_A + mv_B$$
$$3v_A + v_B = 3u - (1)$$

Restitution:

$$\frac{v_B - v_A}{u - 0} = 1$$
$$v_B - v_A = u - (2)$$

$$(1) + 3(2): 4v_B = 6u, v_B = 1.5u, \qquad v_A = 1.5u - u = 0.5u$$

Collision 2: Between B and the wall

Restitution: $\frac{w_B - 0}{0 - 1.5u} = \frac{1}{2} \rightarrow w_B = -0.75\text{u}$

$$t_1 = \frac{a}{1.5u} = \frac{2a}{3u}$$

When A and B collide again, $s_A + s_B = a$.

$$0.5u(t + t_1) + 0.75ut = a$$
$$1.25ut + \frac{a}{3} = a$$
$$t = \frac{2a}{3(1.25u)} = \frac{8a}{15u}$$

Distance from the wall, s_B

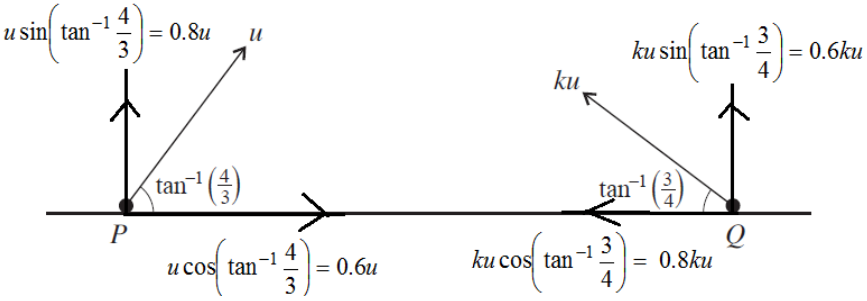
$$= 0.75u\left(\frac{8a}{15u}\right) = 0.4a$$

Example 5: S2008 P2 Q4

Two identical smooth balls, P and Q, are projected simultaneously towards each other from two points on horizontal ground. P is projected with speed u at an angle $\tan^{-1}\left(\frac{4}{3}\right)$ to the horizontal and Q is projected with speed ku at an angle $\tan^{-1}\left(\frac{3}{4}\right)$ to the horizontal (see diagram). The balls collide when they are moving horizontally. It may be assumed that there is no air resistance. Find the value of k. [5]

The coefficient of restitution between the balls is e. Find, in terms of e, u and g, the distance between the points where the balls first hit the ground. [5]

Resolving the velocities into horizontal and vertical components:



Balls collide when they move horizontally:

$$s_y \text{ of } P = s_y \text{ of } Q$$

$$0.8ut - \frac{1}{2}gt^2 = 0.6kut - \frac{1}{2}gt^2$$

$$0.8ut = 0.6kut \rightarrow k = \frac{4}{3}$$

Time to reach ground, t:

$$0.8u = 0 + gt \rightarrow t = \frac{4u}{5g}$$

$$v_{xP} - v_{xQ} = e(u_{xP} - u_{xQ})$$

$$= e\left(0.6u + 0.8\left(\frac{4}{3}\right)u\right) = \frac{5}{3}eu$$

$$\text{Distance} = |v_{xP} - v_{xQ}|t$$

$$= \left(\frac{5}{3}eu\right)\left(\frac{4u}{5g}\right) = \frac{4eu^2}{3g}$$

10.3 Oblique Collisions

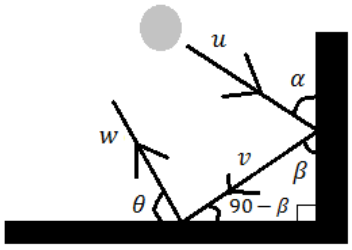
- LO: use conservation of linear momentum and Newton's experimental law to solve problems that may be modelled as the oblique impact of a smooth sphere with a fixed surface.

When a smooth ball strikes a flat surface,

1. The component of its velocity parallel to the surface is unchanged.
2. Newton's law of impact applies to the component of the velocity of the ball perpendicular to the surface before and after impact.
3. The component of its velocity perpendicular to the surface is multiplied by $-e$.
4. The impulse on the ball acts along the line through the centre of the ball perpendicular to the surface, with its magnitude equal to the change of the momentum perpendicular to the surface.

Example 1: W2004 P2 Q2

A snooker ball is moving with speed u towards a point P on one cushion of a snooker table. It rebounds from P and strikes a second cushion at a point Q, from which it rebounds with speed w. The cushions are smooth and perpendicular to each other, and the coefficient of restitution at each impact is e. Show that $w = eu$. [6]



Collision 1:

Consider velocities parallel to P: $u\cos\alpha = v\cos\beta$ $-(1)$

Consider velocities perpendicular to P: $v\sin\beta = eu\sin\alpha$ $-(2)$

Collision 2:

Consider velocities parallel to Q: $v\cos(90-\beta) = w\cos\theta$

$w\cos\theta = eu\sin\alpha$ $-(3)$

Consider velocities perpendicular to Q: $ev\sin(90-\beta) = w\sin\theta$

$w\sin\theta = ev\cos\beta = eu\cos\alpha$ $-(4)$

$(3)^2 + (4)^2: w^2(sin^2\theta + cos^2\theta) = eu(sin^2\alpha + cos^2\alpha)$

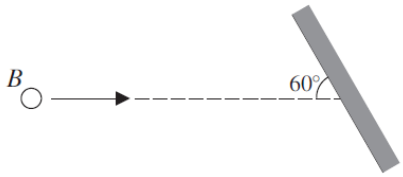
$w^2 = e^2u^2 \rightarrow w = eu$ (shown)

Example 2: S2009 P2 Q5

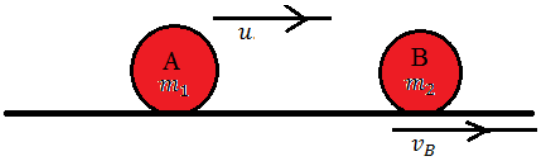
Two spheres A and B, of equal radius, have masses m_1 and m_2 respectively. They lie at rest on a smooth horizontal plane. Sphere A is projected directly towards sphere B with speed u and, as a result of the collision, A is brought to rest. Show that

(i) the speed of B immediately after the collision cannot exceed u, [2]

(ii) $m_1 \leqslant m_2$. [2]



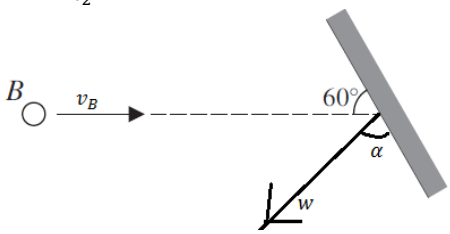
After the collision, B hits a smooth vertical wall which is at an angle of 60° to the direction of motion of B (see diagram). In the impact with the wall B loses $\frac{2}{3}$ of its kinetic energy. Find the coefficient of restitution between B and the wall and show that the direction of motion of B turns through 90°. [8]



(i) $e = \frac{v_B - v_A}{u_A - u_B} \rightarrow v_B = eu$. Since $0 \leq e \leq 1$, $v_B \leq u$.

(ii) Conservation of momentum: $m_1 u = m_2 v_B \rightarrow \frac{v_B}{u} = \frac{m_1}{m_2}$

Since $v_B \leq u, \frac{v_B}{u} \leq 1. \frac{m_1}{m_2} \leq 1, m_1 \leq m_2$.



Consider velocities parallel to the wall: $v_B \cos 60 = w \cos \alpha \quad -(1)$

Consider velocities perpendicular to the wall: $e v_B \sin 60 = w \sin \alpha \quad -(2)$

Loss in KE is $\frac{2}{3}$ of initial KE: $\frac{1}{3}\left(\frac{1}{2}mv_B^2\right) = \frac{1}{2}mw^2 \rightarrow v_B^2 = 3w^2$

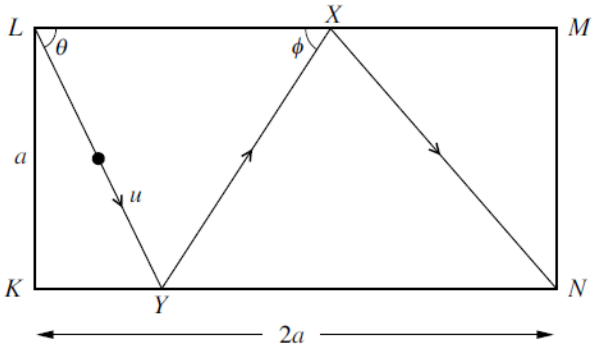
$$(1)^2 + (2)^2: w^2(sin^2\alpha + cos^2\alpha) = v_B^2\left(\frac{1}{4} + \frac{3}{4}e^2\right)$$

$$4w^2 = 3w^2(1 + 3e^2) \rightarrow 1 + 3e^2 = \frac{4}{3}, e^2 = \frac{1}{9}. \text{ Since } e \geq 0, e = \frac{1}{3}.$$

$$\sqrt{3}w\left(\frac{1}{2}\right) = w \cos \alpha \rightarrow \cos \alpha = \frac{\sqrt{3}}{2}, \alpha = 30^o$$

Change in direction = $180^o - 30^o - 60^o = 90^o$ (shown)

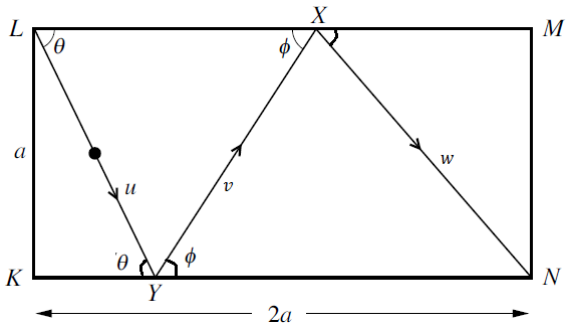
Example 3: S2010 P23 Q5



A rectangular pool table $KLMN$ has $KL = a$ and $KN = 2a$. A ball lies at rest on the table just outside the pocket at L and is projected along the table with speed u in a direction making an angle θ with the edge LM. The ball hits the edge KN at Y, rebounds to hit the edge LM at X and then rebounds into the pocket at N. Angle LXY is denoted by ϕ (see diagram). The coefficient of restitution between the ball and an edge is $\frac{3}{4}$, and all resistances to motion may be neglected. Show that $\tan\phi = \frac{3}{4}\tan\theta$. [3]

Show that $XM = (2 - \frac{7}{3}\cot\theta)a$, and find the value of θ. [6]

Find the speed with which the ball reaches N, giving the answer in the form ku, where k is correct to 3 significant figures. [3]



For collision at Y:

Consider velocities parallel to KL: $u\cos\theta = v\cos\phi \quad -(1)$

Consider velocities parallel to KN: $v\sin\phi = \frac{3}{4}u\sin\theta \quad -(2)$

$\frac{(2)}{(1)}$: $\tan\theta = \frac{3}{4}\tan\phi$ (shown)

$\tan\theta = \frac{a}{LP}, LP = a\cot\theta,$

$\tan\phi = \frac{a}{PX}, PX = a\cot\phi = \frac{4}{3}a\cot\theta$

$XM = 2a - LX$

$= 2a - a\cot\theta - \frac{4}{3}a\cot\theta$

$= \left(2 - \frac{7}{3}\cot\theta\right)a$ (SHOWN)

For collision at X:

Consider velocities parallel to KL: $v\cos\phi = w\cos\beta = u\cos\theta \quad -(3)$

Consider velocities parallel to KN: $w\sin\beta = \frac{3}{4}v\sin\phi = \frac{9}{16}u\sin\theta \quad -(4)$

$\frac{(4)}{(3)}$: $\tan\beta = \frac{MN}{XM} = \frac{9}{16}\tan\theta$

$\frac{16}{9}a\cot\theta = a\left(2 - \frac{7}{3}\cot\theta\right)$

$\frac{37}{9}\cot\theta = 2, \theta = 64.1^o$ to 3 sig. figures.

$(3)^2 + (4)^2$:

$w^2 = (u\cos 64.1)^2 + \left(\frac{9}{16}u\sin 64.1\right)^2$

$w = 0.669u$

CHAPTER 11: CIRCULAR MOTION

11.1 Pre-requisite Knowledge: Refer to Section 10.1 A and 10.1 B (Especially projectiles)

11.2 Pre-requisite Knowledge: Motion in a Horizontal Circle (Paper 5 Review)

- A particle P moving round a circle of radius r with angular speed ω has tangential speed $v = r\omega$. Similarly,

$s = r\theta$, s is linear displacement and θ is angular displacement.
$a = r\alpha$, a is linear acceleration and α is angular acceleration.

- When angular speed of motion is constant, acceleration of the object, $a = \frac{v^2}{r} = \omega^2 r$. This is known as centripetal acceleration as it is always directed to the centre of the circle.
- Centripetal force, $F = ma = m\omega^2 r = \frac{mv^2}{r}$. It is the force that keeps the particle in circular motion and is the resultant force (something has to provide the force).
- When $F_{net} < \frac{mv^2}{r}$, circular motion breaks down.

Example 1: S2004 Mathematics P5 Q7

One end of a light inextensible string of length 0.15 m is attached to a fixed point which is above a smooth horizontal surface. A particle of mass 0.5 kg is attached to the other end of the string. The particle moves with constant speed v m s^{-1} in a horizontal circle, with the string taut and making an angle of $\theta°$ with the downward vertical.

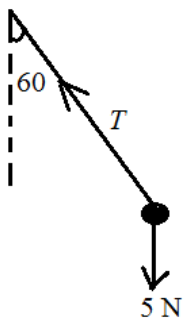
(i) Given that $\theta = 60$ and that the particle is not in contact with the surface, find v. [5]

(ii) Given instead that $\theta = 45$ and $v = 0.9$, and that the particle is in contact with the surface, find

(a) the tension in the string, [2]

(b) the force exerted by the surface on the particle. [3]

(i)



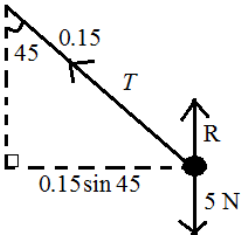
Resolving vertically:

$$T\cos 60 = 0.5(10)$$
$$T = 10N$$

Resolving horizontally:

$$T\sin 60 = \frac{0.5(v^2)}{0.15\sin 60}$$
$$v^2 = 20(0.15)(0.75)$$
$$v = 1.5ms^{-1}$$

(ii) (a)



$$T\sin 45 = \frac{0.5(0.9)^2}{0.15\sin 45}$$
$$T = 5.4N$$

(b)

$$T\cos 45 + R = 0.5(10)$$
$$R = 1.18N$$

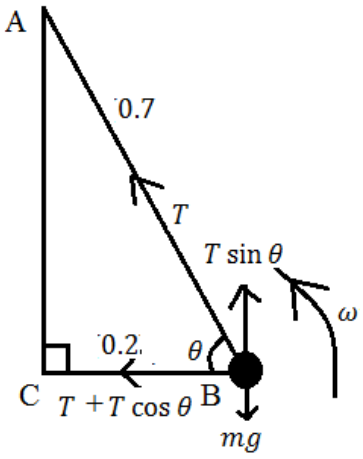
Example 2: W2011 Mathematics P51 Q6

A smooth bead B of mass 0.3 kg is threaded on a light inextensible string of length 0.9 m. One end of the string is attached to a fixed point A, and the other end of the string is attached to a fixed point C which is vertically below A. The tension in the string is T N, and the bead rotates with angular speed ω rad s^{-1} in a horizontal circle about the vertical axis through A and C.

(i) Given that B moves in a circle with centre C and radius 0.2 m, calculate ω, and hence find the kinetic energy of B. [5]

(ii) Given instead that angle $ABC = 90°$, and that AB makes an angle $\tan^{-1}\left(\frac{1}{2}\right)$ with the vertical, calculate T and ω. [6]

(i) Free Body Diagram:



$$\sin\theta = \frac{\sqrt{0.45}}{0.7}, \qquad \cos\theta = \frac{0.2}{0.7} = \frac{2}{7}$$

Resolving vertically: $T\sin\theta = mg$

$$T = \frac{3(0.7)}{\sqrt{0.45}} \approx 3.13\ N$$

Resolving horizontally:

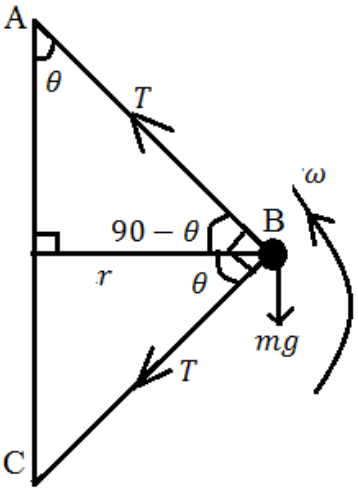
$$T + T\cos\theta = mr\omega^2$$

$$\omega^2 = \frac{3.13\left(\frac{9}{7}\right)}{0.3(0.2)} = 67.07 \rightarrow \omega = 8.19 \text{ rad s}^{-1}$$

$$v = r\omega = 1.64\ m\ s^{-1}$$

$$\text{KE} = \frac{1}{2}(0.3)(1.64)^2 = 0.402 \text{ J}$$

(ii) Free Body Diagram:



$$\tan\theta = \frac{1}{2}, \sin\theta = \frac{1}{\sqrt{5}}, \cos\theta = \frac{2}{\sqrt{5}}$$

Resolving vertically,

$$T\sin\theta + mg = T\cos\theta$$

$$T\left(\frac{2}{\sqrt{5}} - \frac{1}{\sqrt{5}}\right) = 0.3(10)$$

$$T = 3\sqrt{5} \approx 6.71\ N$$

$$AB = \frac{r}{\sin\theta} = \sqrt{5}r, \quad BC = \frac{r}{\cos\theta} = \frac{1}{2}\sqrt{5}r$$

Since $AB + BC = 0.9, r = \frac{1.8}{3\sqrt{5}}$.

Resolving horizontally,

$$T\sin\theta + T\cos\theta = mr\omega^2$$

$$\omega^2 = \frac{3\sqrt{5}}{0.3(1.8)}\left(\frac{3}{\sqrt{5}}\right)(3\sqrt{5}) = 50\sqrt{5}$$

$$\omega = 10.6 \text{ rad s}^{-1}$$

11.3 Motion in a Vertical Circle

- LO: use formulae for the radial and transverse components of acceleration for a particle moving in a circle with variable speed;
- LO: solve problems which can be modelled as a particle, or a pair of connected particles, moving without loss of energy in a vertical circle (including finding the tension in a string or a normal contact force, locating points at which these are zero, and conditions for complete circular motion).

- ❖ Angular displacement: θ, angular velocity, $\omega = \dot{\theta}$, angular acceleration, $\alpha = \dot{\omega} = \ddot{\theta}$.
- ❖ A particle that moves in a vertical circle has variable speed ($v = r\omega$).
- ❖ The acceleration of that particle will have two components:

A transverse / tangential component, $a_t = r\dot{\omega}$
A radial / normal component, $a_r = r\omega^2 = \frac{v^2}{r}$ (centripetal acceleration)
$a = \sqrt{a_t^2 + a_r^2}$, as usual.

- ❖ In a horizontal circle, since speed is constant, $\dot{\omega} = 0$, hence acceleration is only towards the centre of circle.
- ❖ In a vertical circle, as an object moves upwards, it gains potential energy and loses kinetic energy; hence the speed of the particle will differ.
- ❖ To be able to complete a circle for a:

String/Cable, T > 0, $V_{top} > 0$
Sphere (Particle on inside), N > 0, $V_{top} > 0$

- ❖ A particle will not complete a vertical circle if:

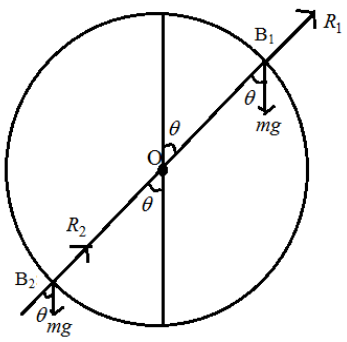
$v = 0$ at any point of the circle (particle will 'oscillate')
$N = 0$ or $T = 0$ at any point (particle will lose contact and travel as a projectile)

Problem Solving Strategies:

1. Draw a free body diagram.
2. Take F = *ma* for normal and tangential directions
3. Use conversation of energy for two points of the circle.
4. Solve simultaneous equations based on the conditions given in the question.
5. Fix conditions to complete a vertical circle, to lose tension, to lose contact etc.
6. Use projectile equations when the particle loses contact.

Example 1: W2008 P2 Q2

A small bead B of mass m is threaded on a smooth wire fixed in a vertical plane. The wire forms a circle of radius a and centre O. The highest point of the circle is A. The bead is slightly displaced from rest at A. When angle $AOB = \theta$, where $\theta < \cos^{-1}\left(\frac{2}{3}\right)$, the force exerted on the bead by the wire has magnitude R_1. When angle $AOB = \pi + \theta$, the force exerted on the bead by the wire has magnitude R_2. Show that $R_2 - R_1 = 4mg$. [8]



Using conservation of energy:

$$\frac{1}{2}mv_{B1}^2 = mg(a - a\cos\theta)$$

$$v_{B1}^2 = 2ag(1-\cos\theta)$$

$$\frac{1}{2}mv_{B2}^2 = \frac{1}{2}mv_{B1}^2 + 2mga\cos\theta$$

$$= mga(1-\cos\theta+2\cos\theta)$$

$$v_{B2}^2 = 2ga(1+\cos\theta)$$

Resolve into the circle at B_1:

$$\frac{mv_{B1}^2}{a} = mg\cos\theta - R_1$$

$$R_1 = mg\cos\theta - 2mg(1-\cos\theta)$$

$$= mg(3\cos\theta - 2)$$

Resolve into the circle at B_2:

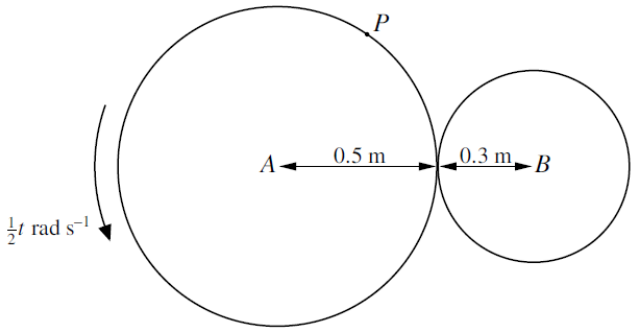
$$\frac{mv_{B2}^2}{a} = R_2 - mg\cos\theta$$

$$R_2 = 2mg(1+\cos\theta) + mg\cos\theta$$

$$= mg(3\cos\theta + 2)$$

(2) – (1): $R_2 - R_1 = mg(3\cos\theta + 2 - 3\cos\theta + 2) = 4mg$ (SHOWN)

Example 2: S2010 P21 Q4



Two coplanar discs, of radii 0.5 m and 0.3 m, rotate about their centres A and B respectively, where AB = 0.8 m. At time t seconds the angular speed of the larger disc is $\frac{1}{2}t$ rad s^{-1} (see diagram). There is no slipping at the point of contact. For the instant when t = 2, find

(i) the angular speed of the smaller disc, [2]

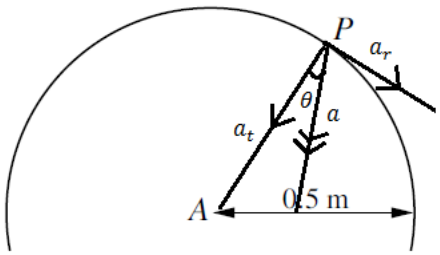
(ii) the magnitude of the acceleration of a point P on the circumference of the larger disc, and the angle between the direction of this acceleration and PA. [7]

(i) Linear Speed of A = Linear Speed of B

$$0.5\omega_A = 0.3\omega_B$$

$$\omega_B = \frac{5}{3}\omega_A = \frac{5}{3}\left(\frac{1}{2}\right)(2) = \frac{5}{3}rad\ s^{-1}$$

(ii)



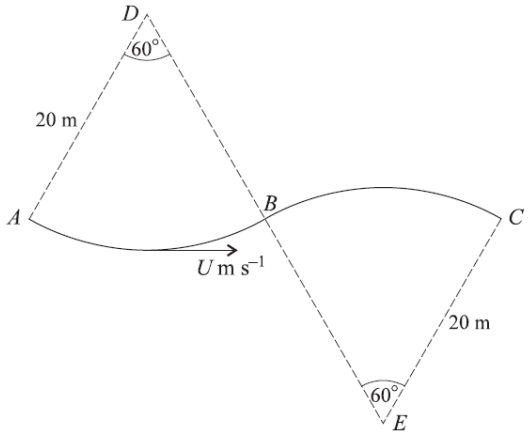
$$a_t = 0.5\frac{d}{dt}\left(\frac{1}{2}t\right) = 0.25ms^{-2}$$

$$a_r = \frac{[0.5(0.5)(2)]^2}{0.5} = 0.5ms^{-2}$$

$$a = \sqrt{a_t^2 + a_r^2} = \sqrt{\frac{5}{16}} = \frac{1}{4}\sqrt{5}$$

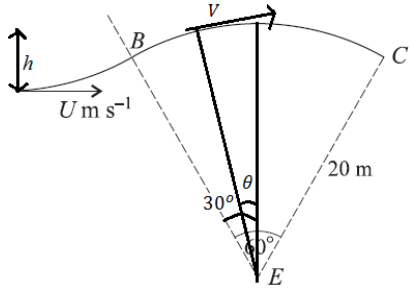
$$\theta = \tan^{-1}\left(\frac{0.25}{0.5}\right) = 26.6^o$$

Example 3: W2002 P2 Q4



A toboggan travels along the path ABC shown in the diagram. The path lies in a vertical plane, and consists of two circular arcs AB and BC. The arcs are of radius 20 m, and subtend angles of 60° at their centres, D and E respectively. The line ABC is horizontal, and there is no friction between the toboggan and the snow. Air resistance is negligible, and the toboggan may be treated as a particle. The speed of the toboggan at its lowest point is U m s^{-1}. Find the range of values of U for which the toboggan will reach C without losing contact with the snow. [10]

Find also the value of U for which the toboggan will leave the snow at B, and travel as a projectile until it lands again at C. [4]



$$h = 20 - 20\cos 30 + 20\cos\theta - 20\cos 30$$
$$= 20\left(1 + \cos\theta - \sqrt{3}\right)$$

Using conservation of energy:

$$\frac{1}{2}mV^2 = \frac{1}{2}mU^2 - mgh$$
$$V^2 = U^2 - 2gh = U^2 - 400\left(1 + \cos\theta - \sqrt{3}\right)$$

The toboggan will not reach C in 2 situations:

Case 1: U is too low where $U < U_{min}$, at $U_{min}, v = 0$ when $\theta = 0$. ('oscillates')

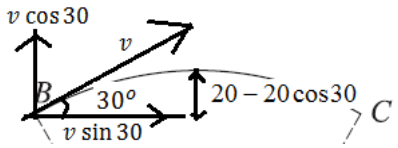
$$U_{min} = 400\left(1 + \cos 0 - \sqrt{3}\right) \approx 10.4 ms^{-1}$$

Case 2: U is too high where $U > U_{max}$, at $U_{max}, N = 0$ when $\theta = 30^o$. (loses contact)

Resolving into the circle at BE:

$$mg\cos\theta - N = \frac{mV^2}{20}$$
$$N = 0, V^2 = 200\cos 30$$
$$U_{max}^2 = 100\sqrt{3} + 400\left(1 + \cos 30 - \sqrt{3}\right)$$
$$U_{max} = 15.1 ms^{-1}$$

Thus to complete a circle, $U_{min} < U < U_{max} \rightarrow 10.4 < U < 15.1$.



$$s_x = 2(20\sin 30)$$
$$v\cos 30\, t = 20$$
$$vt = \frac{20}{\cos 30} \rightarrow t = \frac{40}{\sqrt{3}v} - (1)$$
$$s_y = 0 \rightarrow v\sin 30\, t - \frac{1}{2}gt^2 = 0$$

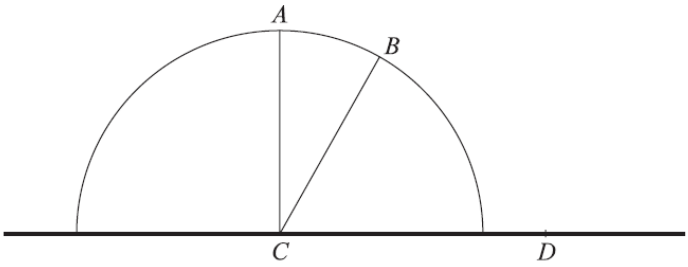
Since $t \neq 0$, $v = 10t - (2)$.

(1) into (2):

$$v = \frac{400}{\sqrt{3}v} \rightarrow v^2 = \frac{400}{\sqrt{3}}$$

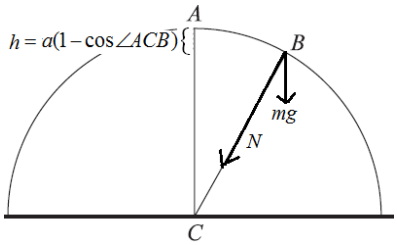
$$\frac{1}{2}mv^2 = \frac{1}{2}mU^2 - mg(20 - 20\cos 30)$$
$$U^2 = v^2 + 20(20 - 20\cos 30)$$
$$= \frac{400}{\sqrt{3}} + 200(2 - \sqrt{3}) \Rightarrow U = 16.9ms^{-1}$$

Example 4: S2004 P2 Q5



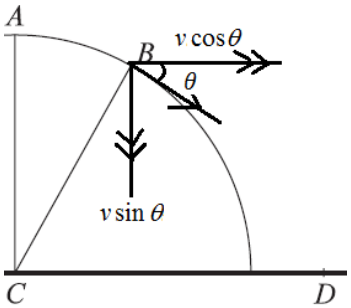
A smooth semicircular cylinder of radius a is fixed on a horizontal table with its axis in the plane of the table. C is a point on the axis and CA is a vertical radius (see diagram). A particle is projected horizontally from A, with speed $\sqrt{(0.4ga)}$, in a direction perpendicular to the axis. Show that the particle leaves the surface with speed $\sqrt{(0.8ga)}$ at the point B where $\cos ACB = 0.8$. [7]

Determine the distance from C of the point D at which the particle hits the table, assuming that it moves freely under gravity from B. [7]



Resolve into the circle at B:

$$mg\cos\angle ACB - N = \frac{mv^2}{a}$$



$\cos\theta = 0.8, \sin\theta = 0.6$

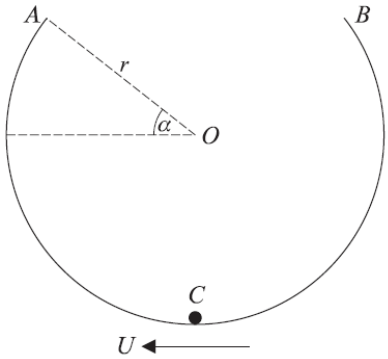
$u = \sqrt{0.4ga}$. Using conservation of energy:

$$\frac{1}{2}mu^2 + mg(a - a\cos\angle ACB) = \frac{1}{2}mv^2$$
$$v^2 = 0.4ga + 2ga(1 - \cos\angle ACB)$$
$$= ga(2.4 - 2\cos\angle ACB)$$

When particle leaves B: $N = 0, v^2 = ga\cos\angle ACB$

$$2.4 - 2\cos\angle ACB = \cos\angle ACB$$
$$\cos\angle ACB = 0.8 \text{ (shown)}$$
$$v^2 = ga(0.8) \rightarrow v = \sqrt{0.8ga} \text{ (shown)}$$
$$s_x = a\cos\theta \rightarrow 0.8a = v\sin\theta t + \frac{1}{2}gt^2$$
$$0.5gt^2 + 0.6t\sqrt{0.8ga} - 0.8a = 0$$
$$t = \frac{-0.6\sqrt{0.8ga} \pm \sqrt{0.36(0.8ga) + 4(0.5g)(0.8a)}}{g}$$
$$= \frac{-0.6\sqrt{0.8a} + \sqrt{1.888a}}{\sqrt{g}}$$
$$s_y = vt\cos\theta$$
$$= 0.8\sqrt{0.8ga}\left(\frac{-0.6\sqrt{0.8a} + \sqrt{1.888a}}{\sqrt{g}}\right) = 0.5992a$$
$$CD = s_y + a\sin\theta$$
$$= 0.5992a + 0.6a \approx 1.20a \text{ (3s.f.)}$$

Example 5: S2005 P2 Q11 EITHER



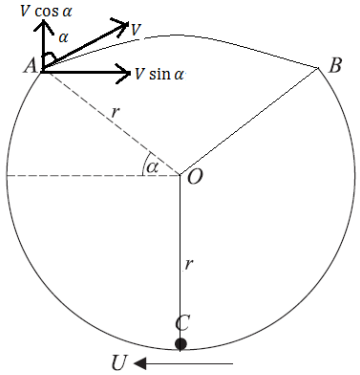
A smooth metal strip with ends A and B is bent into a major arc of a circle with centre O and radius r. The strip is fixed in a vertical plane with AB horizontal and at a height $r\sin\alpha$ above O, where $0 < \alpha < \frac{1}{2}\pi$ (see diagram). A particle is projected from the lowest point C with speed U so that it travels along the inside of the strip, reaching A with speed V. The particle then moves under gravity, and arrives at B. Air resistance may be neglected. Show that

(i) $V^2 = gr\,\text{cosec}\,\alpha$, [4]

(ii) $U^2 = gr(\text{cosec}\,\alpha + 2 + 2\sin\alpha)$. [3]

The mass of the particle is m and the normal reaction between the strip and the particle is R. Show that the greatest value of R is $mg(\text{cosec}\,\alpha + 3 + 2\sin\alpha)$, and find the least value of R. [7]

(i)



$s_x = 2r\cos\alpha$

$V\sin\alpha\,(t) = 2r\cos\alpha - (1)$

$s_y = 0, Vt\sin\alpha - \frac{1}{2}gt^2 = 0$

$2V\cos\alpha = gt, t = \frac{V\cos\alpha}{g} - (2)$

(1) into (2): $\frac{2V^2\sin\alpha\cos\alpha}{g} = 2r\cos\alpha$

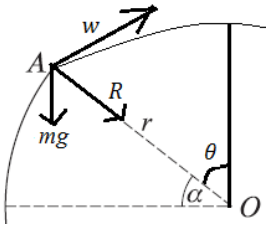
$V^2 = \frac{gr}{\sin\alpha} = gr\csc\alpha$ (SHOWN)

(ii) Using conservation of energy:

$$\frac{1}{2}mU^2 = \frac{1}{2}mV^2 + mgr(1+\sin\alpha)$$

$$U^2 = V^2 + 2gr(1+\sin\alpha)$$

$$= gr(\csc\alpha + 2 + 2\sin\alpha)$$



Resolve into the circle, $mg\cos\theta + R = \frac{mw^2}{r}$.

R_{max} occurs when $\cos\theta = -1, \theta = \pi$ (at C)

$$R_{\max} = \frac{mU^2}{r} + mg$$

$$= mg(3 + 2\sin\alpha + \csc\alpha) \text{ (SHOWN)}$$

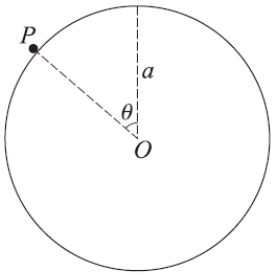
R_{min} occurs when $\cos\theta = \sin\alpha, \theta = \frac{1}{2}\pi - \alpha$ (at A)

Note: $\cos\theta = 1$ is not taken because particle performs projectile motion at $\theta = 0$.

$$R_{\min} = \frac{mV^2}{r} - mg\sin\alpha = mg(\csc\alpha - \sin\alpha)$$

Example 6: S2007 P2 Q11 EITHER

❖ First part of this problem involves knowledge about Chapter 14 or some calculus.



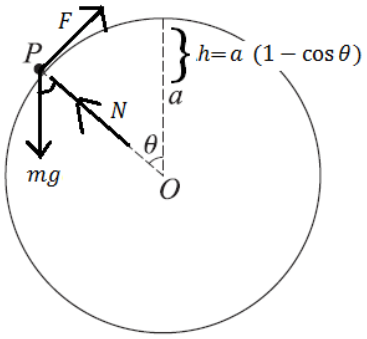
A uniform disc has centre O, radius a and mass $2m$. It is free to rotate in a vertical plane about a horizontal axis through O. A particle P of mass m is placed on the highest point of the rough edge of the disc and the system is slightly disturbed so that OP begins to rotate with the particle in contact with the edge. In the subsequent motion OP makes an angle θ with the upward vertical (see diagram). For the motion while the particle remains in contact with the disc without slipping, find $a\ddot{\theta}$ and $a\dot{\theta}^2$ in terms of g and θ. [6]

Show that if the particle remains in contact with the disc then it begins to slip when

$$4\mu\cos\theta - \sin\theta = 2\mu,$$

where μ is the coefficient of friction between the particle and the edge of the disc. [5]

Show that, however large the value of μ, the particle cannot lose contact with the disc before it starts to slip. [3]



$I = \frac{1}{2}(2m)a^2 + ma^2 = 2ma^2$. Using $\tau = I\ddot{\theta}$:

$$mga\sin\theta = 2ma^2\ddot{\theta} \rightarrow a\ddot{\theta} = \frac{1}{2}g\sin\theta$$

Using conservation of energy:

$$\frac{1}{2}I\dot{\theta}^2 = mg\Delta h$$
$$ma^2\dot{\theta}^2 = mga(1-\cos\theta)$$
$$a\dot{\theta}^2 = g(1-\cos\theta)$$

Alternately, differentiating / integrating one of the equations yield the other one.

Resolving tangentially,

$$mg\sin\theta - F = ma\ddot{\theta}$$
$$F = \frac{1}{2}mg\sin\theta$$

Resolving radially,

$$mg\cos\theta - N = ma\dot{\theta}^2$$
$$N = mg(2\cos\theta - 1)$$

When particle is about to slip, $F = \mu N$.

$$\frac{1}{2}mg\sin\theta = \mu mg(2\cos\theta - 1)$$
$$4\mu\cos\theta - 2\mu = \sin\theta \ \ (shown)$$

$$4\mu\cos\theta - \sin\theta = 2\mu \rightarrow \cos(\theta+\alpha) = \frac{2\mu}{\sqrt{1+16\mu^2}}, \text{where } \tan\alpha = \frac{1}{4\mu}$$

When μ is large, let $\mu = 1$, $\cos(\theta + 14.0^o) = \frac{2}{\sqrt{17}} \rightarrow \theta = 46.9^o$

Hence the particle will slip before $\theta = 46.9^o$.

When the particle loses contact, $N = 0$: $2\cos\theta - 1 = 0 \rightarrow \theta = 60^o$

It will lose contact when $\theta = 60^o$, hence it cannot lose contact before it starts to slip.

CHAPTER 12: EQUILIBRIUM OF A RIGID BODY UNDER COPLANAR FORCES

12.1 Pre-Requisite Knowledge: Paper 4

A) Newton's Laws of Motion

- Newton's 1st Law of Motion: Every object remains in a state of rest or uniform motion in a straight line unless forces act on it to change that state.
- Newton's 2nd Law of Motion: The net force on an object is equal to the product if the mass of the object and its acceleration in the given direction. ($F = ma$)
- Newton's 3rd Law of Motion: If an object A exerts a force on object B, then B exerts a force on A of the same magnitude in the opposite direction.

B) Forces and Equilibrium

- $F_x = F\cos\theta$, $F_y = F\sin\theta$(Resolving horizontally and vertically)
- The resultant of two forces with magnitudes P and Q in perpendicular forces has magnitude $\sqrt{P^2 + Q^2}$ which makes an angle of $\tan^{-1}\left(\frac{Q}{P}\right)$with the direction of the force P.
- When a system is in equilibrium, its resultant force in any direction is equal to zero. That object will either stay at rest (static equilibrium) or move with constant velocity (dynamic equilibrium). Hence, to solve equilibrium problems, take the equations $F_x = 0$ and $F_y = 0$.

12.2 Pre-Requisite Knowledge: Paper 5 (Equilibrium of a Rigid Body)

A) Moment of a Force

- LO: calculate the moment of a force about a point, in two dimensional situations only (understanding of the vector nature of moments is not required);
- LO: recall the principle that if a rigid body is in equilibrium under the action of coplanar forces then the vector sum of the forces is zero and the sum of the moments of the forces about any point is zero, and the converse of this;

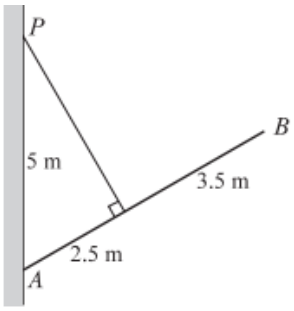
- The rigid body model: an object is assumed to stay the same shape when forces are applied to it (no stretching, squeezing or any deformation).
- The light rod model: weight of components can be neglected (as they are too small to be significant compared to other objects in the system)
- Turning effects of a force is caused by a moment.
- Moment of a force about a point is the product of the magnitude of the force and the distance of its line of action from the point.

Moment = Force × Perpendicular distance. SI unit: Newton metre (Nm)

- For an object to be in equilibrium, clockwise moments = anticlockwise moments. This equation is known as taking moments about a particular point.

- ❖ Principle of moments: if a rigid object is in equilibrium, the sum of clockwise moments about any point must be equal to the sum of anticlockwise moments about that point.
- ❖ General strategy: An unknown or unwanted force can be removed by taking moments about a point on its line of action.

Example: S2003 Mathematics P5 Q3

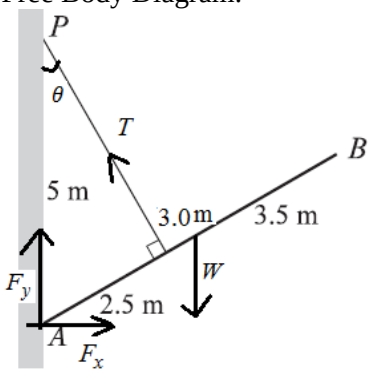


A uniform beam AB has length 6 m and mass 45 kg. One end of a light inextensible rope is attached to the beam at the point 2.5 m from A. The other end of the rope is attached to a fixed point P on a vertical wall. The beam is in equilibrium with A in contact with the wall at a point 5 m below P. The rope is taut and at right angles to AB (see diagram). Find

(i) the tension in the rope, [4]

(ii) the horizontal and vertical components of the force exerted by the wall on the beam at A. [3]

Free Body Diagram:



(i) $$\sin\theta = \frac{2.5}{5.0} \rightarrow \theta = 30^o$$

Taking moments of the beam about A,

$$T(2.5) = (3.0)(45)(10)\cos\theta$$

$$T = \frac{1350}{2.5}\cos 30 = 270\sqrt{3} \approx 468N$$

(ii) Resolving horizontally:

$$F_x = T\sin 30 = 135\sqrt{3} \approx 244N$$

Resolving vertically:

$$F_y + T\cos 30 = W$$

$$F_y = 450 - 405 = 45N$$

B) Centre of Mass

- LO: understand and use the result that the effect of gravity on a rigid body is equivalent to a single force acting at the centre of mass of the body, and identify the centre of mass by considerations of symmetry in suitable cases.

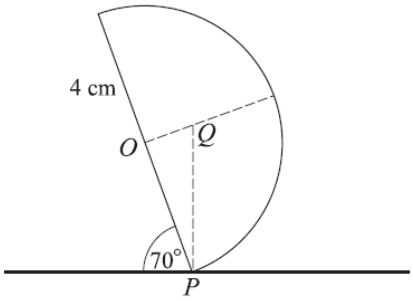
- ❖ A rigid body always have a point, known as a centre of mass, where the mass of the whole body can be assumed to act on.
- ❖ Uniform model: object made of the same material with same density all the way through. For any uniform rigid object, the centre of mass is the centre of symmetry.
- ❖ For a one-dimensional rod: centre of mass, $\bar{x} = \frac{m_1x_1 + m_2x_2 + \cdots + m_nx_n}{M}$.
- ❖ It is basically equivalent to taking a moment equation about one end of the rod.

- For a two-dimensional object, a similar formula is used for both x and y coordinates: $\bar{x} = \frac{m_1x_1+m_2x_2+\cdots+m_nx_n}{M}$, $\bar{y} = \frac{m_1y+m_2y_2+\cdots+m_ny_n}{M}$.
- Centre of mass of some common objects (quoted from MF10):

Triangular lamina: $\frac{2}{3}$ along median from vertex
Solid hemisphere of radius r: $\frac{3}{8}r$ from centre
Hemispherical shell of radius r: $\frac{1}{2}r$ from centre
Circular arc of radius r and angle 2α: $\frac{r\sin\alpha}{\alpha}$ from centre
Circular sector of radius r and angle 2α: $\frac{2r\sin\alpha}{3\alpha}$ from centre
Solid cone or pyramid of height h: $\frac{3}{4}h$ from vertex

- Exact methods of calculating centre of mass are required for M2 only. In FM2 centre of mass calculation only involves the formulae given above (this will be re-used in Chapter 13: Rotation of a Rigid Body)

Example 1: S2003 Mathematics P5 Q2



A uniform solid hemisphere, with centre O and radius 4 cm, is held so that a point P of its rim is in contact with a horizontal surface. The plane face of the hemisphere makes an angle of 70° with the horizontal. Q is the point on the axis of symmetry of the hemisphere which is vertically above P. The diagram shows the vertical cross-section of the hemisphere which contains O, P and Q.

(i) Determine whether or not the centre of mass of the hemisphere is between O and Q. [3]

The hemisphere is now released.

(ii) State whether or not the hemisphere falls on to its plane face, giving a reason for your answer. [2]

(i) Let the centre of mass of the hemisphere be at point G.

$$OQ = 4\tan 20 = 1.456\ cm$$

Centre of mass of a hemisphere is $\frac{3}{8}r$ from centre:

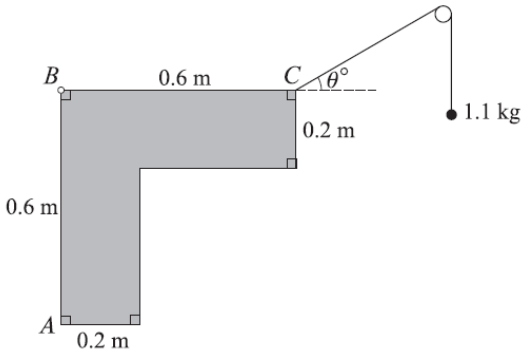
$$OG = \frac{3}{8}(4) = 1.5\ cm$$

Thus the centre of mass is not between O and Q.

(ii) When moments of the hemisphere are taken about P, a resultant clockwise moment is present (due to weight of the hemisphere).

Hence the hemisphere does not fall on its plane face.

Example 2: S2006 Mathematics P5 Q5



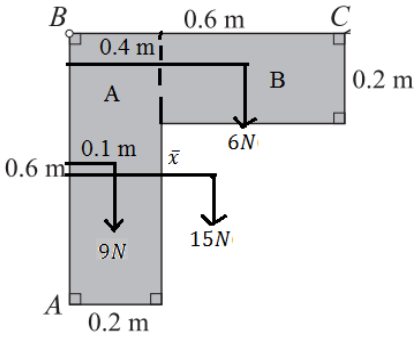
A uniform lamina of weight 15 N has dimensions as shown in the diagram.

(i) Show that the distance of the centre of mass of the lamina from AB is 0.22 m. [4]

The lamina is freely hinged at B to a fixed point. One end of a light inextensible string is attached to the lamina at C. The string passes over a fixed smooth pulley and a particle of mass 1.1 kg is attached to the other end of the string. The lamina is in equilibrium with BC horizontal. The string is taut and makes an angle of $\theta°$ with the horizontal at C, and the particle hangs freely below the pulley (see diagram).

(ii) Find the value of θ. [3]

(i) Free Body Diagram:



Weight of A $= \frac{0.6\times0.2}{1.0\times0.2} \times 15N = 9N$

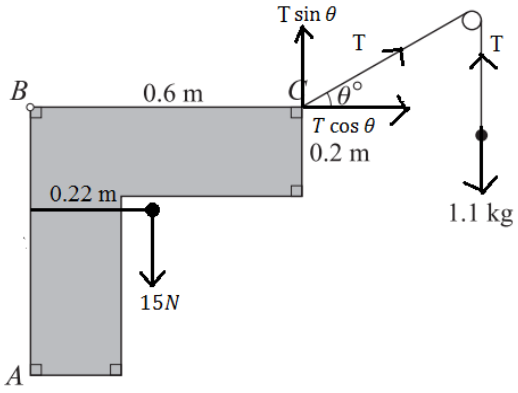
Weight of B $= \frac{0.4\times0.2}{1.0\times0.2} \times 15N = 6N$

Taking moments of the lamina along AB,

$$15N(\bar{x}) = 9N(0.1m) + 6N(0.4m)$$

$$\bar{x} = \frac{0.9m + 2.4m}{15} = 0.22\ m$$

(ii) Free Body Diagram:



Resolving vertically for particle:

$$T = 1.1(10) = 11\ N$$

Taking moments of the system about B,

$$1.1(10)(0.6\sin\theta) = 15(0.22)$$

$$\sin\theta = \frac{15(0.22)}{11(0.6)} = 0.5$$

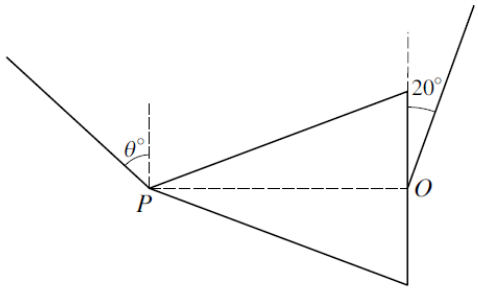
$$\theta = 30^o$$

C) Equilibrium of a Rigid Body

- LO: solve problems involving the equilibrium of a single rigid body under the action of coplanar forces (problems set will not involve complicated trigonometry).

❖ To solve equilibrium points, take moment equations and resolve in both directions. Solve simultaneous equation if necessary to find unknown forces.
❖ For equilibrium, $F \leq \mu R$. $F = \mu R$ occurs during limiting equilibrium.
❖ Equilibrium of an object can be broken by sliding (if there is a resultant force) or toppling (if there is a resultant moment). Both processes can occur at the same time.
❖ In a rigid body, line of action of forces (especially normal forces) can be obtained by taking moments to find the distance from a particular point where the force acts.
❖ Angle of friction: Sometime frictional forces act at an angle (this is often neglected due to resolving the components in other directions).
❖ We know that $F = W\sin\theta$ and $R = W\cos\theta$. Hence $\tan\theta = \frac{F}{R}$.
❖ For equilibrium, $F \leq \mu R$, thus $\tan\theta \leq \mu, \theta \leq \tan^{-1}\mu$.
❖ Note: Problems involving sliding and toppling is only involved in M2 and not FM2.

Example 1: W2009 Mathematics P52 Q6

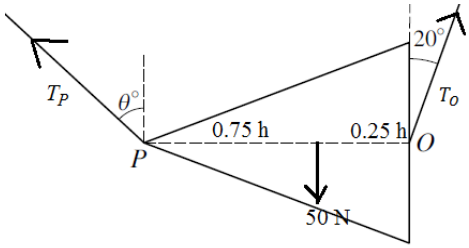


P is the vertex of a uniform solid cone of mass 5 kg, and O is the centre of its base. Strings are attached to the cone at P and at O. The cone hangs in equilibrium with PO horizontal and the strings taut. The strings attached at P and O make angles of $\theta°$ and 20°, respectively, with the vertical (see diagram, which shows a cross-section).

(i) By taking moments about P for the cone, find the tension in the string attached at O. [4]

(ii) Find the value of θ and the tension in the string attached at P. [6]

(i) Free Body Diagram:



Taking moments of the cone about P,

$$T_O \cos 20\ (h) = 5(10)\left(\frac{3}{4}h\right)$$

$$T_O = \frac{150}{4\cos 20} \approx 39.9\ N$$

(ii) Resolving horizontally,

$$T_P \sin\theta = T_O \sin 20 = 13.65N - (1)$$

Resolving vertically,

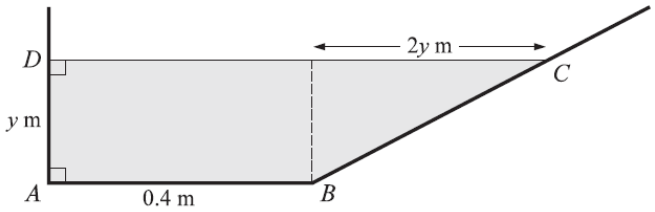
$$T_P \cos\theta + T_O \cos 20 = 50$$

$$T_P \cos\theta = 12.497\ N - (2)$$

$(1)^2 + (2)^2$: $\quad T_P = \sqrt{13.65^2 + 12.497^2} = 18.5\ N$

$$\frac{(1)}{(2)}: \tan\theta = \frac{13.65}{12.497} \rightarrow \theta = 47.5^o$$

Example 2: W2004 Mathematics P5 Q7



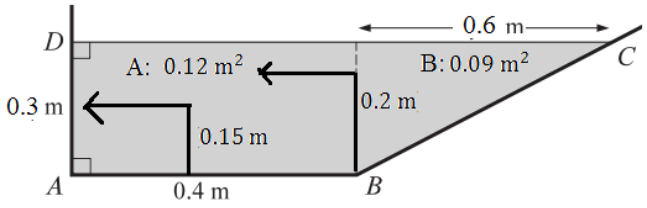
A light container has a vertical cross-section in the form of a trapezium. The container rests on a horizontal surface. Grain is poured into the container to a depth of y m. As shown in the diagram, the cross-section $ABCD$ of the grain is such that $AB = 0.4$ m and $DC = (0.4 + 2y)$ m.

(i) When $y = 0.3$, find the vertical height of the centre of mass of the grain above the base of the container. [5]

(ii) Find the value of y for which the container is about to topple. [5]

(i) Area of A = $0.4\,y$, Area of B = $0.5(y)(2y) = y^2$

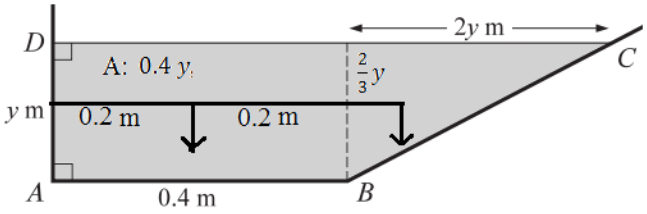
When $y = 0.3\,m$, Free Body Diagram:



Taking moments along AB, $0.12\,(0.15\text{ m}) + 0.09(0.2\text{ m}) = 0.21\,\bar{y}$

$$\bar{y} = \frac{0.12(0.15) + 0.09(0.2)}{0.21} = \frac{6}{35}\text{m}$$

(ii) Free Body Diagram:



Taking moments along AD, $0.2\,(0.4y) + \left(0.4 + \frac{2}{3}y\right)\,(y^2) = \bar{x}(0.4y + y^2)$

$$\bar{x} = \frac{0.08 + 0.4y + \frac{2}{3}y^2}{0.4 + y}$$

When the container is about to topple, $\bar{x} = 0.4\,m$:

$$0.08 + 0.4y + \frac{2}{3}y^2 = 0.16 + 0.4y$$

$$\frac{2}{3}y^2 = 0.08 \rightarrow y^2 = \frac{3}{25} \rightarrow y = \frac{\sqrt{3}}{5}$$

12.3 System of Rigid Bodies

- LO: use Newton's third law in situations involving the contact of rigid bodies in equilibrium.

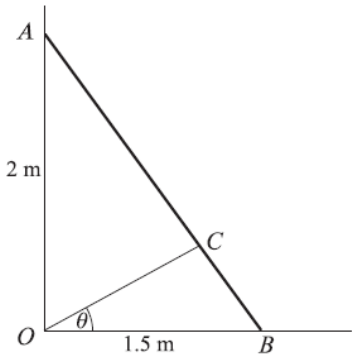
A) Revisiting Newton's 3rd Law

- ❖ When two objects are joined to each other through a pin joint, the forces exerted on them by the pin have equal magnitudes and opposite directions.

B) Problem Solving Strategies:

1. Draw a free body diagram (either for the whole or part of the system). Label and mark everything so that it is not confusing (ensure that direction of normal and frictional forces are correct)
2. Take equilibrium force equations: vertical forces = 0, horizontal forces = 0
3. Take moment equations: clockwise moments = anticlockwise moments
4. For system of objects, resolve for the whole system and take moments by the contact point as normal and frictional forces will not be taken into consideration this way.
5. Solve simultaneous equations if necessary to find unknown forces.
6. For equilibrium, $F \leq \mu N$. For questions which have different coefficient of frictions at different parts, take the higher value of μ to ensure equilibrium for the whole system.
7. To find forces and angles which they act, resolve horizontally and vertically then combine them using the method shown in Section 12.1 B.

Example 1: S2005 P2 Q4



A uniform ladder AB, of weight W N and length 2.5 m, rests against a smooth vertical wall OA with its foot on smooth horizontal ground OB. The ladder is in a vertical plane perpendicular to the wall. It is kept in position with $OA = 2$ m and $OB = 1.5$ m by a light rope OC joining O to a point C on the ladder, such that angle $COB = \theta$. Show that the tension T N in the rope is given by

$$T = \frac{3W}{8\cos\theta - 6\sin\theta}. \quad [7]$$

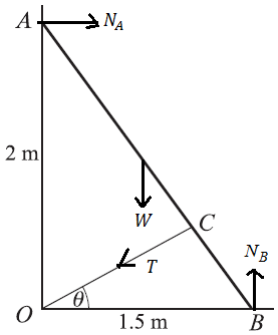
Deduce that $\tan\theta < \frac{4}{3}$. [2]

Find T in terms of W in the case when the rope is perpendicular to the ladder. [2]

Free Body Diagram:

Resolve horizontally: $N_A = T\cos\theta$

Resolve vertically: $N_B = W + T\sin\theta$

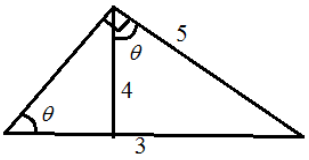


Taking moments of the system about O,

$$N_B(1.5) = W(0.75) + N_A(2)$$
$$1.5W + 1.5T\sin\theta = 0.75W + 2T\cos\theta$$
$$T(2\cos\theta - 1.5\sin\theta) = 0.75W$$
$$T = \frac{0.75W}{2\cos\theta - 1.5\sin\theta}$$
$$= \frac{3W}{8\cos\theta - 6\sin\theta}$$

If $T = 0$, $F_x = N_A$, hence the system cannot be in equilibrium. Thus T > 0:

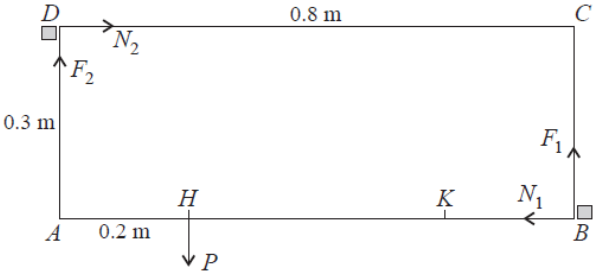
$$\frac{3W}{8\cos\theta - 6\sin\theta} > 0 \Rightarrow 8\cos\theta - 6\sin\theta > 0$$
$$6\sin\theta < 8\cos\theta, \tan\theta < \frac{4}{3}$$



If $OC \perp AB$, $\tan\theta = \frac{3}{4}, \sin\theta = \frac{3}{5}, \cos\theta = \frac{4}{5}$

$$T = \frac{3W}{8(0.8) - 6(0.6)} = \frac{15}{14}W$$

Example 2: W2003 P2 Q11 EITHER



A rectangular drawer $ABCD$, where AD = 0.3 m, DC = 0.8 m, can move on a smooth horizontal surface if its two handles at H and K are pulled horizontally perpendicular to its front edge AB. An attempt is made to open the drawer by pulling on only one handle H, where AH = 0.2 m. In this case, frictional forces F_1 and F_2, at B and D respectively, from the sides of the drawer frame impede opening (see diagram). The normal reactions at B and D are N_1 and N_2 respectively.

(i) Assuming that the coefficient of friction at B is 0.8, and that the contact at D is smooth, show that the drawer cannot be opened by a force P applied to H. [4]

(ii) Assume now that $\frac{F_1}{N_1} = \frac{F_2}{N_2}$ at all times, and that the coefficients of friction at B and D are each equal to μ.

(a) Find the value that μ must exceed if the drawer cannot be opened by a force P applied to H. [4]

(b) An attempt is made to open the drawer by applying at H, in addition to the force P, a horizontal force Q along the front edge. The reactions of the sides of the drawer frame on the drawer still act at B and D. Given that μ = 0.8, and that the drawer is on the point of moving, find the direction of Q and the ratio of the magnitude of Q to the magnitude of P. [6]

(i) Taking moments about C: $P(0.6) = N_1(0.3) \rightarrow P = 0.5N_1$
For equilibrium, $F_1 \leq \mu N_1$. Thus $F_{max} = 0.8N_1$.
Resultant force downwards = $P - F_1$.
Since $P < F_{max}$, resultant force < 0, hence the drawer cannot be opened.

(ii)(a) Resolving horizontally: $N_1 = N_2 = N$. Thus it follows that $F_1 = F_2 = F$.
Taking moments about H: $F(0.2) + N(0.3) = F(0.6) \rightarrow F = 0.75N$
For equilibrium, $F \leq \mu N$. Thus μ must exceed 0.75.

(ii)(b) Since drawer is at limiting equilibrium,.$F_1 = 0.8N_1, F_2 = 0.8N_2$.
Resolving horizontally, $Q + N_2 = N_1 - (1)$
Resolving vertically, $F_1 + F_2 = P - (2)$
Taking moments about H: $F_1(0.6) = F_2(0.2) + N_2(0.3)$

$$0.8N_1(0.6) = 0.8N_2(0.2) + N_2(0.3)$$

$$0.48N_1 = 0.46N_2 \Rightarrow N_2 = \frac{24}{23}N_1.$$

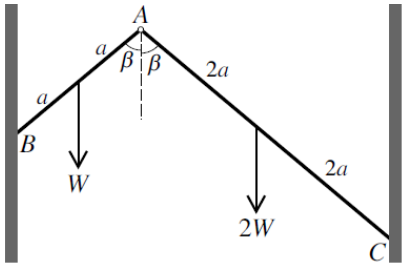
Substitute into (1): $$Q = N_1 - \frac{24}{23}N_1 = -\frac{1}{23}N_1$$

Substitute into (2): $$P = 0.8N_1 + 0.8\left(\frac{24}{23}N_1\right) = \frac{188}{115}N_1$$

Since Q and N_1 are in opposite directions, direction of Q is towards A.

$$\left|\frac{Q}{P}\right| = \frac{1}{23}N_1 \div \frac{188}{115}N_1 = \frac{5}{188}$$

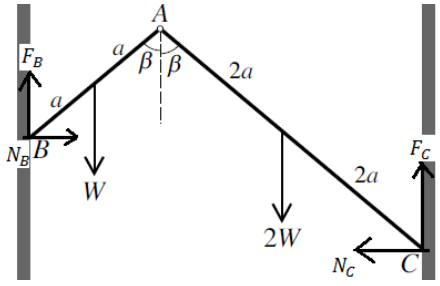
Example 3: S2010 P23 Q11 EITHER



Two uniform rods AB and AC have lengths $2a$ and $4a$ and weights W and $2W$ respectively. They are freely hinged together at A and rest in equilibrium in a vertical plane with B and C in contact with two rough parallel vertical walls. The plane containing the rods is perpendicular to the walls. The rods AB and AC each make an angle β with the vertical (see diagram). Show that the magnitude of the frictional force acting on AB at B is $\frac{5}{4}W$. [7]

Given that the coefficient of friction at B and at C is μ, find the set of possible values of μ in terms of β. [7]

Free Body Diagram:



Taking moments about B for rod AB: $F_B(2a\sin\beta) = N_B(2a\cos\beta) + Wa\sin\beta$
$2N_B\cos\beta = 2F_B\sin\beta - W\sin\beta - (1)$

Taking moments about C for the system: $F_B(6a\sin\beta) + N_B(2a\cos\beta) = 2W(2a\sin\beta) + W(5a\sin\beta)$
$2N_B\cos\beta = 9W\sin\beta - 6F_B\sin\beta - (2)$

$(1) = (2)$: $9W\sin\beta - 6F_B\sin\beta = 2F_B\sin\beta - W\sin\beta$

$$8F_B = 10W \rightarrow F_B = \frac{5}{4}W \text{ (shown)}$$

$F_B = \frac{5}{4}W$ into (1), $\quad 2N_B\cos\beta = \frac{5}{2}W\sin\beta - W\sin\beta \rightarrow N_B = \frac{3}{4}W\tan\beta$

Resolving vertically for the system: $F_B + F_C = 3W \rightarrow F_C = \frac{7}{4}W$

Resolving horizontally for the system: $N_B = N_C = \frac{3}{4}W\tan\beta$

For equilibrium at B: $F_B \leq \mu N_B$

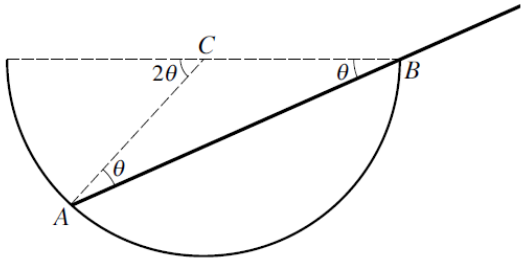
$\frac{3}{4}W\tan\beta\,\mu \geq \frac{5}{4}W \rightarrow \mu \geq \frac{5}{3}\cot\beta$

For equilibrium at C: $F_C \leq \mu N_C$

$\frac{3}{4}W\tan\beta\,\mu \geq \frac{7}{4}W \rightarrow \mu \geq \frac{7}{3}\cot\beta$

Hence, for equilibrium of the whole system, $\mu \geq \frac{7}{3}\cot\beta$.

Example 4: W2010 P2 Q4



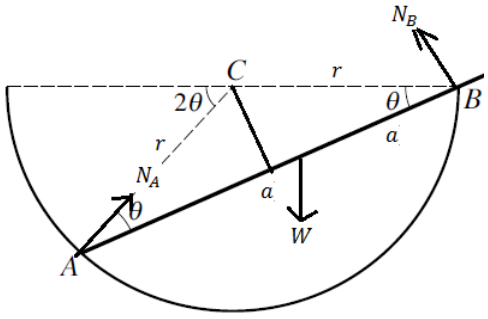
A hemispherical bowl of radius r is fixed with its rim horizontal. A thin uniform rod rests in equilibrium on the rim of the bowl with one end resting on the inner surface of the bowl at A, as shown in the diagram. The rod has length $2a$ and weight W. The point of contact between the rod and the rim is B, and the rim has centre C. The rod is in a vertical plane containing C. The rod is inclined at θ to the horizontal and the line AC is inclined at 2θ to the horizontal. The contacts at A and B are smooth. In any order, show that

(i) the contact force acting on the rod at A has magnitude $W\tan\theta$,

(ii) the contact force acting on the rod at B has magnitude $\frac{W\cos 2\theta}{\cos\theta}$,

(iii) $2r\cos 2\theta = a\cos\theta$. [9]

Free Body Diagram:



Taking moments of the rod about B,

$N_A\sin\theta\,(2r\cos\theta) = W\cos\theta\,(2r\cos\theta - a)$

$N_A(2r\sin\theta) = W(2r\cos\theta - a)$

$N_A = \frac{W(2r\cos\theta - a)}{2r\sin\theta}$

Taking moments of the rod about A:

$W\cos\theta\,(a) = N_B(2r\cos\theta) \rightarrow N_B = \frac{Wa}{2r}$

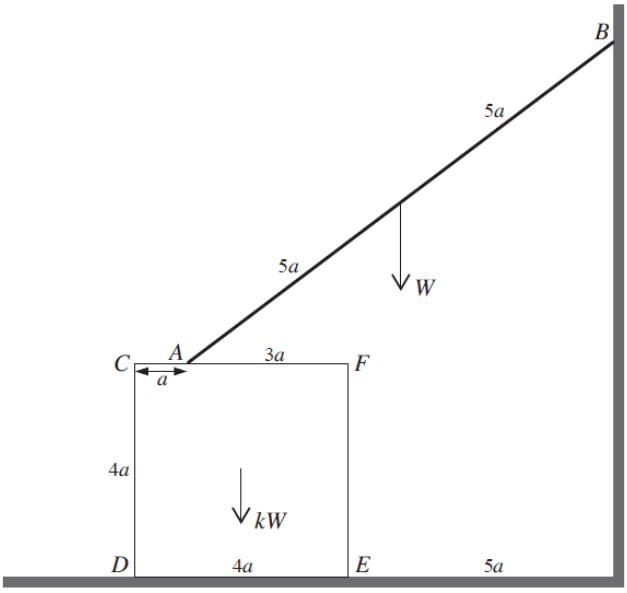
Resolving horizontally for system,

$N_B\sin\theta = N_A\cos 2\theta$

$$\frac{Wa}{2r}\sin\theta = \frac{W(2r\cos\theta - a)}{2r\sin\theta}\cos 2\theta$$
$$a\sin^2\theta = \cos 2\theta\,(2r\cos\theta - a)$$
$$a(\sin^2\theta + \cos 2\theta) = 2r\cos 2\theta\cos\theta$$
$$2r\cos 2\theta\cos\theta = a\cos^2\theta$$
$$2r\cos 2\theta = a\cos\theta \text{ [Result(iii)]}$$
$$N_B = \frac{W}{2}\left(\frac{2\cos 2\theta}{\cos\theta}\right)$$
$$= \frac{W\cos 2\theta}{\cos\theta} \text{ [Result(ii)]}$$

$$N_A = \frac{W(2r\cos\theta - \frac{2r\cos 2\theta}{\cos\theta})}{2r\sin\theta}$$
$$N_A\sin\theta\cos\theta = W(\cos^2\theta - \cos 2\theta)$$
$$= W(\cos^2\theta - 2\cos^2\theta + 1)$$
$$= W(1 - \cos^2\theta)$$
$$N_A = \frac{W\sin^2\theta}{\sin\theta\cos\theta}$$
$$= W\tan\theta \text{ [Result (i)]}$$

Example 5: W2008 P2 Q11 EITHER

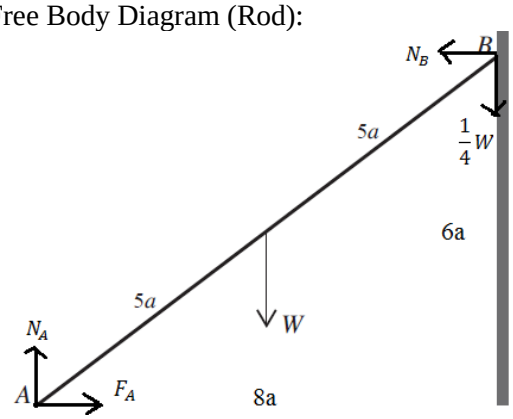


The diagram shows a central cross-section $CDEF$ of a uniform solid cube of weight kW with edges of length $4a$. The cube rests on a rough horizontal floor. One of the vertical faces of the cube is parallel to a smooth vertical wall and at a distance $5a$ from it. A uniform ladder, of length $10a$ and weight W, is represented by AB. The ladder rests in equilibrium with A in contact with the rough top surface of the cube and B in contact with the wall. The distance AC is a and the vertical plane containing AB is perpendicular to the wall. The coefficients of friction between the ladder and the cube, and between the cube and the floor, are both equal to μ. A small dog of weight $\frac{1}{4}W$ climbs the ladder and reaches the top without the ladder sliding or the cube turning about the edge through D. Show that $\mu \geqslant \frac{4}{5}$. [7]

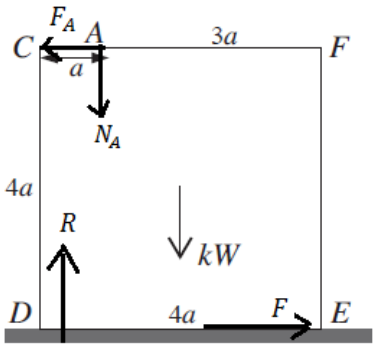
Show that the cube does not slide whatever the value of k. [4]

Find the smallest possible value of k for which equilibrium is not broken. [3]

Free Body Diagram (Rod):



Free Body Diagram (Cube):



Resolve vertically for rod: $N_A = W + \frac{1}{4}W = \frac{5}{4}W$

Taking moments of the rod about A: $N_B(6a) = W(4a) + \frac{1}{4}(8a)W$

$6N_B = 6W \rightarrow N_B = W$

Resolve horizontally for rod: $F_A = F_B = W$

For equilibrium: $F_A \leq \mu N_A$

$W \leq \frac{5}{4}W \rightarrow \mu \geq \frac{4}{5}$ $(shown)$

Resolve horizontally for cube: $F = F_A = W$

Resolve vertically for cube: $R = kW + \frac{5}{4}W$

$$\mu_{min}R = \frac{4}{5}\left(k + \frac{5}{4}\right)W = \left(\frac{4}{5}k + 1\right)W$$

Since $k > 0$: $\mu_{min}R > W$ for all k.

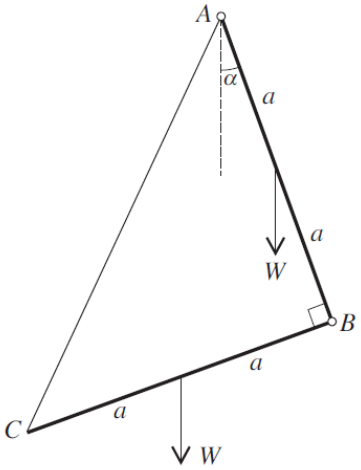
$F = W \leq \mu_{min}R, F < \mu R$ for all k, hence cube does not slide for any value of k.

Taking moments of the cube about D: $F_A(4a) = N_A(a) + k_{min}W(2a)$

$2k_{min}W = 4W - \frac{5}{4}W = \frac{11}{4}W$

$k_{min} = \frac{11}{8}$

Example 6: W2009 P2 Q5

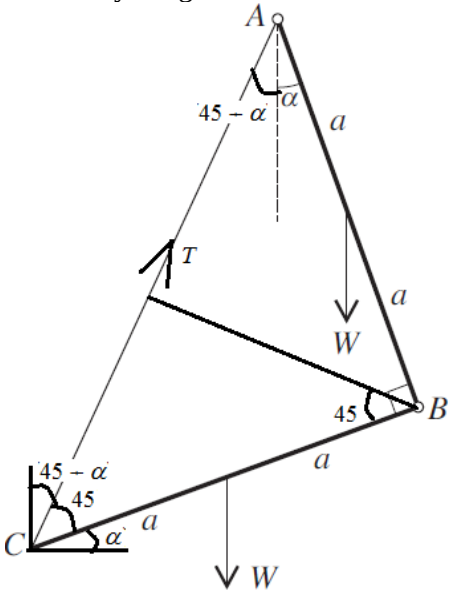


Two uniform rods, AB and BC, each have length $2a$ and weight W. They are smoothly jointed at B, and A is attached to a smooth fixed pivot. A light inextensible string of length $(2\sqrt{2})a$ joins A to C so that angle $ABC = 90°$. The system hangs in equilibrium, with AB making an angle α with the vertical (see diagram). By taking moments about A for the system, or otherwise, show that $\alpha = 18.4°$, correct to the nearest $0.1°$. [3]

Find the tension in the string in the form kW, giving the value of k correct to 3 significant figures. [3]

Find, in terms of W, the magnitude of the force acting on the rod BC at B. [6]

Free Body Diagram:



Taking moments of the system about A:

$$W(a\sin\alpha) = W(a\cos\alpha - 2a\sin\alpha)$$

$$\sin\alpha = \cos\alpha - 2\sin\alpha$$

$$\tan\alpha = \frac{1}{3} \rightarrow \alpha \approx 18.4°$$

Taking moments of rod BC about B:

$$T(2a\cos 45) = W\cos\alpha$$

$$T = \frac{3}{\sqrt{20}}W \approx 0.671W$$

Resolve horizontally for rod BC:

$$F_x = T\cos(45+\alpha) = \frac{3}{10}W$$

Resolve vertically for rod BC:

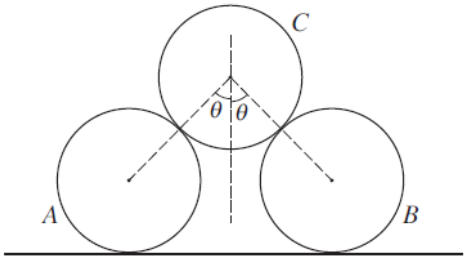
$$F_y + T\sin(45+\alpha) = W$$

$$F_y = W - \frac{3}{5}W = \frac{2}{5}W$$

$$F = \sqrt{F_x^2 + F_y^2}$$

$$= \sqrt{(0.3W)^2 + (0.4W)^2} = \frac{1}{2}W$$

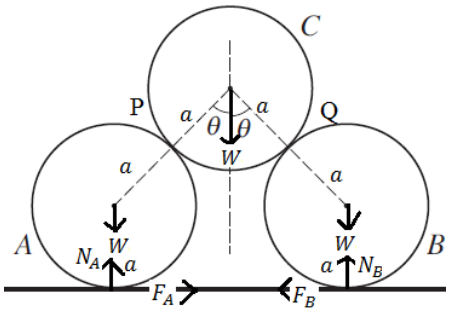
Example 7: W2012 P21 Q10 EITHER



Two identical uniform rough spheres A and B, each of weight W and radius a, are at rest on a rough horizontal plane, and are not in contact with each other. A third identical sphere C rests on A and B with its centre in the same vertical plane as the centres of A and B. The line joining the centres of A and C and the line joining the centres of B and C are each inclined at an angle θ to the vertical (see diagram). The coefficient of friction between each sphere and the plane is μ. The coefficient of friction between C and A, and between C and B, is μ'. The system remains in equilibrium. Show that

$$\mu \geqslant \frac{\sin\theta}{3(1+\cos\theta)} \quad \text{and} \quad \mu' \geqslant \frac{\sin\theta}{1+\cos\theta}. \qquad [14]$$

Free Body Diagram: System



Resolving vertically for system: $N_A + N_B = 3W$

Resolving horizontally for system: $F_A = F_B$

Taking moments about point P:

$$N_A(a\sin\theta) = W(a\sin\theta) + F_A(a + a\cos\theta)$$

Taking moments about point Q:

$$N_B(a\sin\theta) = W(a\sin\theta) + F_B(a + a\cos\theta)$$

$F_A = F_B$, thus $N_A(a\sin\theta) = N_B(a\sin\theta)$

$$N_A = N_B = \frac{3}{2}W$$

$$\frac{3}{2}Wa\sin\theta = Wa\sin\theta + F_A(a + a\cos\theta)$$

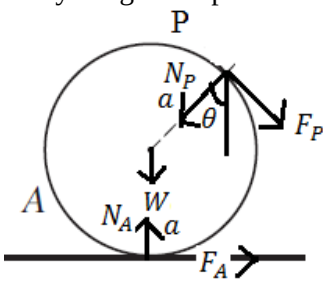
$$F_A = \frac{W\sin\theta}{2(1+\cos\theta)}$$

For equilibrium, $F_A \le \mu N_A$,

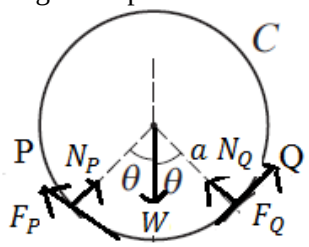
$$\frac{3}{2}\mu W \ge \frac{W\sin\theta}{2(1+\cos\theta)}$$

$$\mu \ge \frac{W\sin\theta}{3(1+\cos\theta)} \ (SHOWN)$$

Free Body Diagram: Sphere A



Free Body Diagram: Sphere C



Taking moments about centre of C, $F_P = F_Q$.

Taking moments about centre of A: $F_P = F_A$.

Resolving horizontally for sphere C:

$$F_Q\cos\theta + N_P\sin\theta = F_P\cos\theta + N_Q\sin\theta$$

$$N_P = N_Q$$

Resolving vertically for sphere C:

$$F_P\sin\theta + F_Q\sin\theta + N_P\cos\theta + N_Q\cos\theta = W$$

$$2F_P\sin\theta + 2N_P\cos\theta = W$$

$$N_P\cos\theta = \frac{1}{2}W - \frac{W\sin^2\theta}{2(1+\cos\theta)}$$

$$= \frac{1}{2}W\left(\frac{1+\cos\theta-\sin^2\theta}{1+\cos\theta}\right)$$

$$= \frac{1}{2}W\left(\frac{\cos^2\theta+\cos\theta}{1+\cos\theta}\right)$$

$$N_P = \frac{1}{2}W$$

For equilibrium, $F_P \le \mu' N_P$

$$\frac{1}{2}\mu' W \ge \frac{W\sin\theta}{2(1+\cos\theta)}$$

$$\mu' \ge \frac{\sin\theta}{1+\cos\theta} \ (SHOWN)$$

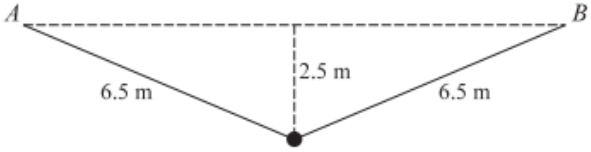
CHAPTER 13: SIMPLE HARMONIC MOTION

Note: Although being listed as Topic 14 in the syllabus, the real Topic 13 (Rotation of a Rigid Body) requires full understanding of this chapter and hence it will be covered first.

13.1 Hooke's Law (Paper 5 Review)

- Force applied is directly proportional to extension of material.
- Elastic model: An object that will return to its original length when external forces applied on it is removed.
- If too much force is applied, material will lose its elasticity (permanent deformation).
- Hooke's law: $T = \frac{\lambda x}{l}$, x = extension, l = natural length.
- Modulus of elasticity, λ: depends on type of material used, its unit is newton (N).
- Elastic potential energy, $E_p = \frac{1}{2}kx^2 = \frac{\lambda x^2}{2l}$.

Example 1: S2003 Mathematics P5 Q7



A light elastic string has natural length 10 m and modulus of elasticity 130 N. The ends of the string are attached to fixed points A and B, which are at the same horizontal level. A small stone is attached to the mid-point of the string and hangs in equilibrium at a point 2.5 m below AB, as shown in the diagram. With the stone in this position the length of the string is 13 m.

(i) Find the tension in the string. [2]

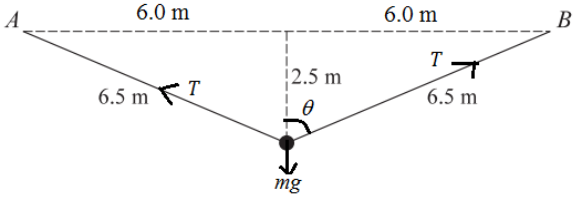
(ii) Show that the mass of the stone is 3 kg. [2]

The stone is now held at rest at a point 8 m vertically below the mid-point of AB.

(iii) Find the elastic potential energy of the string in this position. [3]

(iv) The stone is now released. Find the speed with which it passes through the mid-point of AB. [4]

Free body diagram:

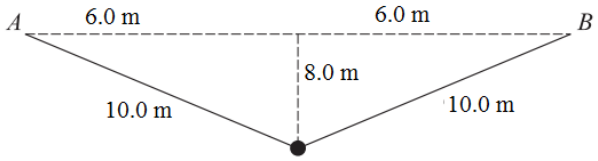


(i) $T = \frac{130(13-10)}{10} = 39N$

(ii) Resolving downwards:

$2T\cos\theta = mg$

$2(39)\left(\frac{2.5}{6.5}\right) = 10m \Rightarrow m = 3kg$



(iii) $E_p = \dfrac{130(20-10)^2}{2(10)}$

$= 650J$

(iv) Using conservation of energy:

$$\text{Loss in EPE} = \text{Gain in KE} + \text{Gain in GPE}$$

$$650 - \frac{130(12-10)^2}{2(10)} = \frac{1}{2}(3)v^2 + 3(10)(8)$$

$$\frac{3}{2}v^2 = 650 - 26 - 240$$

$$v^2 = 256 \rightarrow v = 16\ ms^{-1}$$

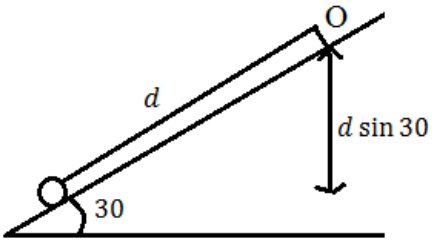
Example 2: W2002 Mathematics P5 Q5

A light elastic string has natural length 2 m and modulus of elasticity 1.5 N. One end of the string is attached to a fixed point O of a smooth plane which is inclined at 30° to the horizontal. The other end of the string is attached to a particle P of mass 0.075 kg. P is released from rest at O. Find

(i) the distance of P from O when P is at its lowest point, [5]

(ii) the acceleration with which P starts to move up the plane immediately after it has reached its lowest point. [4]

(i)



At the lowest point, KE $= 0$.

Using conservation of energy:

$$\text{Gain in EPE} = \text{Loss in GPE}$$

$$0.075\ (10)(d \sin 30) = \frac{1.5(d-2)^2}{2(2)}$$

$$0.75\ d \sin 30 = \frac{0.75}{2}(d-2)^2$$

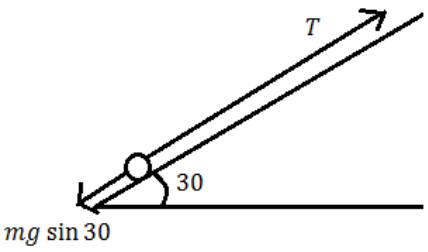
$$(d-2)^2 = d \rightarrow d^2 - 5d + 4 = 0$$

$$(d-1)(d-4) = 0$$

$$d = 1 \text{ or } d = 4$$

Since > 2, $d = 4m$.

(ii)



$$T = \frac{1.5(4-2)}{2} = 1.5\ N$$

Resolving parallel to the plane,

$$T - mg \sin 30 = ma$$

$$0.075a = 1.5 - 0.75 \sin 30$$

$$a = \frac{1.125}{0.075} = 15\ ms^{-2}$$

13.2 Simple Harmonic Motion: Terminologies

- LO: recall a definition of SHM and understand the concepts of period and amplitude;
- LO: use standard SHM formulae in the course of solving problems;

- ❖ Definition (S2003 P2 Q1): SHM is a type of periodic motion where magnitude of the acceleration of the body is directly proportional to its displacement from equilibrium and the direction of the acceleration always acts towards the equilibrium point.
- ❖ Equations of motion: $\ddot{x} = -\omega^2 x$, $x = a\cos(\omega t + \varepsilon)$, $x = a\sin(\omega t + \varepsilon)$
  [This can be solved using the Differential Equations method in Topic 6]
- ❖ $\omega$is the angular frequency, $\omega = 2\pi f = \frac{2\pi}{T}$, its unit is rad s^{-1}.
- ❖ Period, T: time taken to complete one oscillation, $T = \frac{2\pi}{\omega}$.
- ❖ Frequency, f: number of oscillations per unit time, $f = \frac{1}{T}$.
- ❖ Amplitude, a: maximum distance of the body from the equilibrium point.
- ❖ Equations involving velocity:

  - $x = a\sin(\omega t + \varepsilon) \Rightarrow v = \omega a\cos(\omega t + \varepsilon)$
  - $x = a\cos(\omega t + \varepsilon) \Rightarrow v = -\omega a\sin(\omega t + \varepsilon)$
  - $v^2 = \omega^2(a^2 - x^2)$

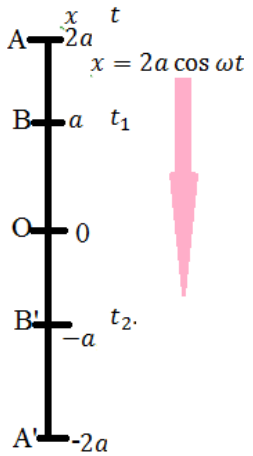
- ❖ For motion equations, the sine version is used if the particle starts at the origin and the cosine version is used if the particle starts at the end points.
- ❖ A negative sign is sometimes added depending on how the particle moves. Generally $x = \pm a\sin\omega t$ or $x = \pm a\cos\omega t$ is used.

Example 1: S2008 P2 Q2

A particle P moves on a straight line AOA' in simple harmonic motion with period T. The centre of the motion is O and P is instantaneously at rest at A and A'. The mid-points of OA and OA' are B and B' respectively.

(i) Find, in terms of T, the time for P to move directly from B to B'. [5]

(ii) Find the ratio of the speed of P at B to its speed at O. [3]



(i) When $x = a$, $t = t_1$.

$2a\cos\omega t_1 = a$

$\cos\omega t_1 = 0.5$

$\omega t_1 = \frac{1}{3}\pi \rightarrow t_1 = \frac{\pi}{3\omega}$

When $x = -a$, $t = t_2$.

$2a\cos\omega t_2 = -a$

$\cos\omega t_2 = -0.5$

$\omega t_2 = \frac{2}{3}\pi \rightarrow t_2 = \frac{2\pi}{3\omega}$

$$Time = t_2 - t_1 = \frac{\pi}{3\left(\frac{2\pi}{T}\right)} = \frac{1}{6}T$$

(ii) At B: $v_B^2 = \omega^2(4a^2 - a^2) = 3a^2\omega^2$

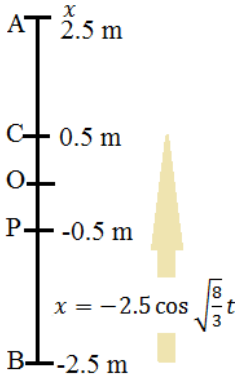
At O: $v_O^2 = \omega^2(4a^2 - 0^2) = 4a^2\omega^2$

$$\frac{v_B}{v_O} = \sqrt{\frac{3a^2\omega^2}{4a^2\omega^2}} = \frac{\sqrt{3}}{2}$$

Example 2: W2007 P2 Q2

A particle P oscillates in simple harmonic motion between the points A and B, where $AB = 5$ m. When P is 2 m from B its speed is 4 m s^{-1}. Find the number of complete oscillations that P makes in one minute. [5]

The point C is between A and B, and $BC = 3$ m. Find the time taken for P to move directly from B to C. [3]



(i) $A = 2.5$, When $x = -0.5m, v = 4ms^{-1}$.

$$16 = \omega^2(2.5^2 - 0.5^2) \rightarrow \omega^2 = \frac{16}{6} = \frac{8}{3}$$

No. of oscillations in a minute	$= \frac{60}{T} = \frac{30\omega}{\pi} = \frac{30}{\pi}\sqrt{\frac{8}{3}} \approx 15.6$

Thus P performs 15 complete oscillations in a minute.

(ii) Equation of motion is $x = -2.5\cos\sqrt{\frac{8}{3}}t$. At C, $x = 0.5m$:

$$0.5 = -2.5\cos\sqrt{\frac{8}{3}}t$$

$$\cos\sqrt{\frac{8}{3}}t = -0.2$$

$$\sqrt{\frac{8}{3}}t = \cos^{-1}(-0.2) = \pi - \cos^{-1}0.2$$

$$t = \sqrt{\frac{3}{8}}\left[\pi - \cos^{-1}0.2\right] = 1.09s$$

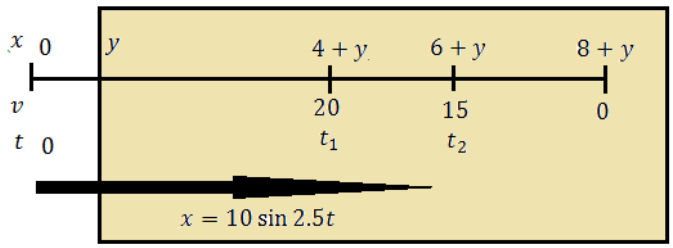
Example 3: W2005 P2 Q4

A point on a screen is moving in a horizontal straight line with simple harmonic motion. The point is off the screen for some of its motion, and the centre of the motion is not clear. The speeds at which the point crosses certain vertical lines, at different distances from the left-hand edge of the screen, are measured. The results are given in the table.

Distance from edge of screen (cm)	4	6	8
Speed of point (cm s^{-1})	20	15	0

Find the position of the centre of the simple harmonic motion, and the time it takes the point to move directly from the 4 cm line to the 6 cm line. [10]

Let the centre of motion be y cm left of the screen. Thus $a = 8 + y$.



When $x = 4 + y, v = 20$.

$$400 = \omega^2[(8+y)^2 - (4+y)^2]$$

$$\omega^2 = \frac{400}{48 + 8y} - (1)$$

When $x = 6 + y, v = 15$.

$$225 = \omega^2[(8+y)^2 - (6+y)^2]$$

$$\omega^2 = \frac{225}{28 + 4y} - (2)$$

$$(1) = (2): \frac{225}{28 + 4y} = \frac{400}{48 + 8y}$$

$$16(7+y) = 9(12+2y) \rightarrow y = 2.$$

Hence centre of motion is 2cm left from the screen.

$$\omega^2 = \frac{225}{36} \rightarrow \omega = 2.5 rad\ s^{-1}$$

Equation of motion is $x = 10\sin 2.5t$.

When $x = 6, t = t_1$.

$10\sin 2.5t_1 = 6 \rightarrow t_1 = 0.4\sin^{-1} 0.6$

When $x = 8, t = t_2$.

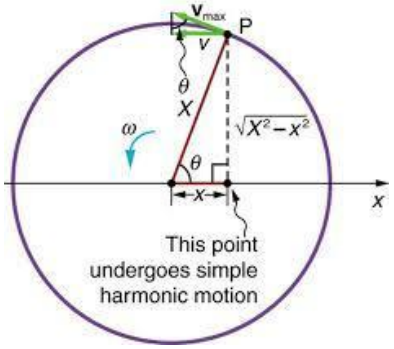
$10\sin 2.5t_2 = 8 \rightarrow t_2 = 0.4\sin^{-1} 0.8$

Time taken is $t_2 - t_1 = 0.4\sin^{-1} 0.8 - 0.4\sin^{-1} 0.6 = 0.114$ seconds, correct to 3 significant figures.

13.3 Oscillations and Approximations

- LO: set up the differential equation of motion in problems leading to SHM, recall and use appropriate forms of solution, and identify the period and amplitude of the motion;
- LO: recognise situations where an exact equation of motion may be approximated by an SHM equation, carry out necessary approximations (e.g. small angle approximations or binomial approximations) and appreciate the conditions necessary for such approximations to be useful.

- ❖ Shifting the origin: The differential equation $\ddot{x} = c - kx = -k\left(x - \frac{c}{k}\right)$ can be reduced to SHM form $\ddot{y} = -ky$ using the substitution $y = x - \frac{c}{k}$ (this is very useful when finding centre of motion)
- ❖ SHM and Circular Motion: SHM along a line can be illustrated as the projection on the line of a point which moves round a circle with its centre of the oscillation, radius the amplitude, with constant angular speed ω, where the period of the oscillation is $T = \frac{2\pi}{\omega}$.



❖ Approximations:

1. Trigonometric approximations: When θ is small, $\sin\theta \approx \theta, \cos\theta \approx 1, \tan\theta \approx \theta$

(Applying the Maclaurin series: ignoring powers of 2 and higher)

2. Binomial expansion approximation (P3): When x is small ($|x|<1$),

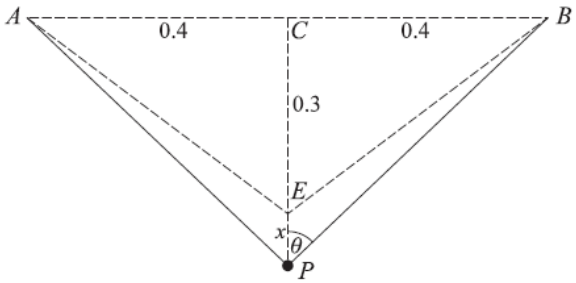
$$(1+x)^n = 1 + \frac{nx}{1!} + \frac{n(n-1)x^2}{2!} + \cdots$$

Problem Solving Strategies:

1. Draw a free body diagram.
2. Take equation of motion $F = m\ddot{x}$ and express the equation in the form of $\ddot{x} = -\omega^2 x$.
3. Use binomial or trigonometric approximations if required.
4. Determine period through the formula $T = \frac{2\pi}{\omega}$.
5. Determine the amplitude of the motion.
6. Select the appropriate equation of motion and solve for time taken, displacement etc.
7. When the string becomes slack consider motion under gravity / projectile motion.

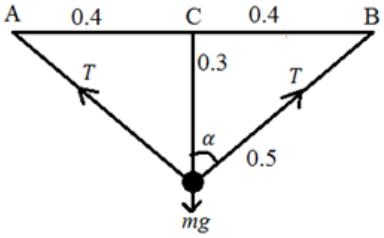
Example 1: W2002 P2 Q10 EITHER

A light elastic string, of natural length 0.8 m and with modulus of elasticity 40 N, is attached at its ends to fixed points, A and B, at the same level and 0.8 m apart. A particle P of mass 1.2 kg is attached to the mid-point of the string. The point E is vertically below the mid-point C of AB, and $CE = 0.3$ m. Show that the particle can remain in equilibrium at E. [4]



The particle is released from rest at a point vertically below E. Its distance below E at time t s after release is x m, where x is small (see diagram). Show that, if x^2 and higher powers of x are neglected, the tension in each of the strings AP and BP is $10(1 + 6x)$ N, and that each string makes an angle θ with the vertical, where $\cos\theta = 0.6 + 1.28x$. [6]

Find the approximate period of small oscillations of P. [4]



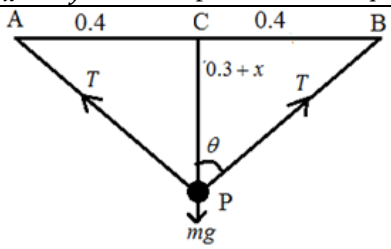
$$T = \frac{40(0.5-0.4)}{0.4} = 10N$$

Resolving horizontally: $F_x = T\sin\alpha - T\sin\alpha = 0$

Resolving vertically:

$$F_y = 2T\cos\alpha - mg = 2(10)\left(\frac{0.3}{0.5}\right) - 1.2(10) = 0$$

Since $F_x = F_y = 0$, the particle is in equilibrium at E.



Pythagoras theorem: $BP^2 = 0.4^2 + (0.3+x)^2$

Neglecting x^2 and higher terms:

$$BP = (0.25+0.6x)^{0.5}$$
$$= \sqrt{0.25}(1+2.4x)^{0.5}$$
$$\approx 0.5(1+1.2x)$$

$$T = \frac{40(BP-0.4)}{0.4}$$
$$= 100(0.1+0.6x)$$
$$= 10(1+6x)N$$

$$\cos\theta = \frac{CP}{BP}$$
$$= (0.3+x)(0.5+0.6x)^{-1}$$
$$= (0.6+2x)(1+1.2x)^{-1}$$
$$\approx (0.6+2x)(1-1.2x)$$
$$= 0.6+1.28x$$

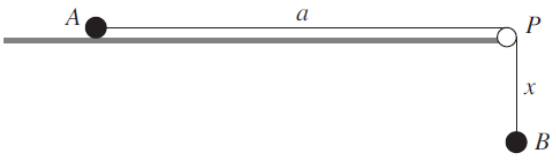
Resolving vertically: $m\ddot{x} = mg - 2T\cos\theta$

$$1.2\ddot{x} = 12 - 20(0.6+1.28x)(1+6x)$$
$$\approx 12 - 20(0.6x+4.88x)$$
$$= -97.6x \rightarrow \ddot{x} = -\frac{244}{3}x$$

Hence the motion is simple harmonic with $\omega = \sqrt{\frac{244}{3}}$.

$$T = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{3}{244}} \approx 0.691s\ (3\ s.f.)$$

Example 2: S2009 P2 Q11 EITHER



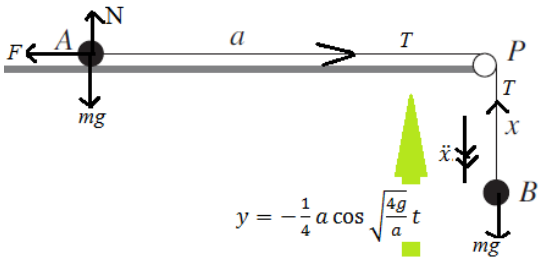
Two particles A and B, of equal mass m, are connected by a light elastic string of natural length a and modulus of elasticity $4mg$. Particle A rests on a rough horizontal table at a distance a from the edge of the table. The string passes over a small smooth pulley P fixed at the edge of the table. At time $t = 0$, B is released from rest at P and falls vertically. At time t, B has fallen a distance x, without A slipping (see diagram). Show that

$$\ddot{x} = -\frac{g}{a}(4x-a). \qquad [3]$$

Deduce that, while A does not slip, B moves in simple harmonic motion and identify the centre of the motion. [3]

Given that the coefficient of friction between A and the table is $\frac{1}{3}$, find the value of x when A starts to slip, and the corresponding value of t, expressing this answer in the form $k\sqrt{\left(\frac{a}{g}\right)}$. Give the value of k correct to 3 decimal places. [8]

Free Body Diagram:



$$T = \frac{4mgx}{a}$$

Resolve vertically at a general point:

$$m\ddot{x} = mg - T$$
$$= mg - \frac{4mgx}{a}$$
$$\ddot{x} = -\frac{g}{a}(4x - a)$$

At equilibrium, $\frac{4mgx}{a} = mg \rightarrow x = \frac{1}{4}a$. Thus centre of motion is $\frac{1}{4}a$.

Let $y = x - \frac{1}{4}a \rightarrow \dot{y} = \dot{x} \rightarrow \ddot{y} = \ddot{x}$. Thus $\ddot{y} = -\frac{4g}{a}y$.

If A does not slip, B moves in simple harmonic motion with $\omega^2 = \frac{4g}{a}$.

Resolve horizontally for A, when particle is about to move:

$$T = F$$
$$\frac{4mgx}{a} = \frac{1}{3}mg$$
$$x = \frac{1}{12}a$$

When $x = \frac{1}{12}a \rightarrow y = -\frac{1}{6}a$

$$-\frac{1}{6}a = -\frac{1}{4}a\cos\sqrt{\frac{4g}{a}}t$$

$$\cos\sqrt{\frac{4g}{a}}t = \frac{2}{3} \rightarrow t = \sqrt{\frac{a}{4g}}\cos^{-1}\left(\frac{2}{3}\right) = 0.421\sqrt{\frac{a}{g}}$$

Thus $k = 0.421$ to 3 decimal places.

Example 3: W2009 P2 Q11 EITHER

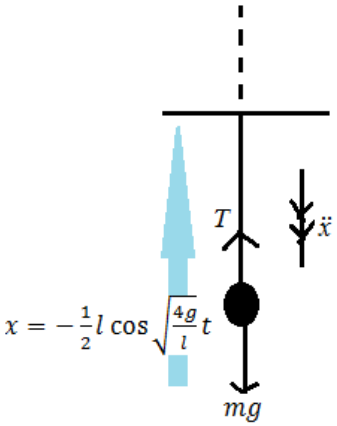
A light elastic string, of natural length l and modulus of elasticity $4mg$, is attached at one end to a fixed point and has a particle P of mass m attached to the other end. When P is hanging in equilibrium under gravity it is given a velocity $\sqrt{(gl)}$ vertically downwards. At time t the downward displacement of P from its equilibrium position is x. Show that, while the string is taut,

$$\ddot{x} = -\frac{4g}{l}x. \qquad [4]$$

Find the speed of P when the length of the string is l. [4]

Show that the time taken for P to move from the lowest point to the highest point of its motion is

$$\left(\frac{\pi}{3} + \frac{\sqrt{3}}{2}\right)\sqrt{\left(\frac{l}{g}\right)}. \qquad [6]$$



At equilibrium, $\frac{4mg(e)}{l} = mg \rightarrow e = \frac{1}{4}l$

Resolve downwards at a general point:

$$m\ddot{x} = mg - T$$
$$= mg - \frac{4mg(x + 0.25l)}{l}$$
$$= mg - mg - \frac{4mgx}{l} \Rightarrow \ddot{x} = -\frac{4gx}{l}$$

$\omega^2 = \frac{4g}{l}$. When $x = 0,\ v = \sqrt{gl}$.

$$gl = \frac{4g}{l}(A^2 - 0^2) \rightarrow A^2 = \frac{1}{4}l$$

When length is l, $x = -\frac{1}{4}l$

$$v^2 = \frac{4g}{l}\left(\frac{1}{4}l^2 - \frac{1}{16}l^2\right)$$
$$= 0.75gl \Rightarrow v = \sqrt{0.75gl}$$

Part 1: Particle moves under SHM

When $x = -\frac{1}{4}l$, $t = t_1$.

$$-\frac{1}{4}l = -\frac{1}{2}l\cos\sqrt{\frac{4g}{l}}t_1$$

$$\cos\sqrt{\frac{4g}{l}}t_1 = \frac{1}{2} \Rightarrow t_1 = \frac{2\pi}{3}\sqrt{\frac{l}{4g}} = \frac{\pi}{3}\sqrt{\frac{l}{g}}$$

Part 2: Particle moves under gravity (after string becomes slack)

When $= 0$, $t = t_2$.

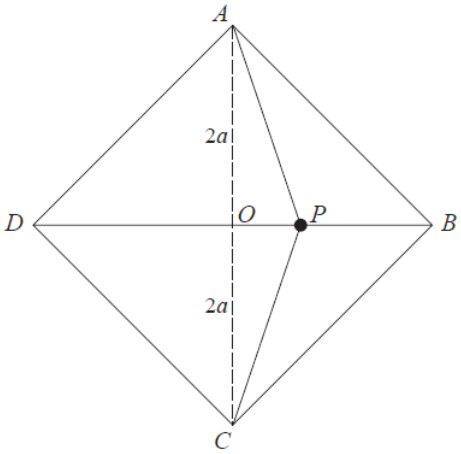
Using $v = u + at$:

$$0 = \sqrt{0.75gl} - gt_2$$

$$t_2 = \sqrt{\frac{3l}{4g}} = \frac{\sqrt{3}}{2}\sqrt{\frac{l}{g}}$$

$$t = t_1 + t_2 = \left(\frac{\pi}{3} + \frac{\sqrt{3}}{2}\right)\sqrt{\frac{l}{g}}(SHOWN)$$

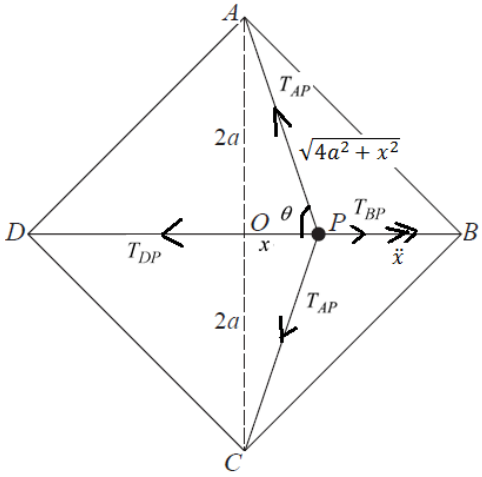
Example 4: W2003 P2 Q4



ABCD is a smooth square horizontal table with centre *O* and diagonals of length $4a$. A particle *P* of mass *m* rests on the table and is connected to *A* and to *C* by two light elastic strings, each of natural length a and modulus of elasticity λ (see diagram). Show that, when *P* is displaced a short distance x from *O* towards *B*, the resultant force towards *O* is $\frac{\lambda x}{a}$, if $\left(\frac{x}{a}\right)^2$ and higher powers of $\frac{x}{a}$ are neglected. [5]

Two additional strings, *PB* and *PD*, are now introduced, so that *P* is connected to all four corners of the table by strings of natural length a and modulus of elasticity λ. The particle is displaced from *O* a short distance towards *B*, and is released from rest. Show that *P* describes approximate simple harmonic motion along *DB*, and find the period of this motion. [6]

Free Body Diagram:



$$T_{AP} = \frac{\lambda(AP - a)}{a} = \frac{\lambda}{a}\left(\sqrt{4a^2 + x^2} - a\right)$$

Resultant force towards O , $F = 2\,T_{AP}\cos\theta$

$$F = \frac{2\lambda}{a}\left(\sqrt{4a^2 + x^2} - a\right)\left(\frac{x}{\sqrt{4a^2 + x^2}}\right)$$

$$= \frac{2\lambda x}{a}\left[1 - a(4a^2)^{-\frac{1}{2}}\left(1 + \frac{x^2}{4a^2}\right)^{-\frac{1}{2}}\right]$$

$$= \frac{2\lambda x}{a}\left[1 - \frac{a}{2a}(1 + \cdots)\right] \approx \frac{\lambda x}{a}$$

Resolving horizontally:

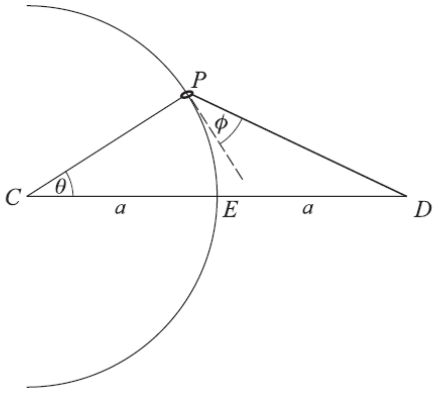
$$m\ddot{x} = T_{BP} - T_{DP} - F$$

$$= \frac{(2a - x - a)\lambda}{a} - \frac{(2a + x - a)\lambda}{a} - \frac{\lambda x}{a}$$

$$= \frac{\lambda}{a}(a - x - a - x - x) = \frac{-3\lambda x}{a}$$

$\ddot{x} = \frac{-3\lambda}{ma}x$. Hence it is a SHM with $\omega^2 = \frac{3\lambda}{ma}$. $T = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{ma}{3\lambda}}$.

Example 5: W2004 P2 Q11 EITHER



A ring P of mass m kg is threaded on a smooth semicircular wire with centre C and radius a m, fixed in a horizontal plane. A light elastic string connects P to D, a fixed point on the axis of the semicircle with $CD = 2a$ m (see diagram). The natural length of the string is less than a m. P is displaced through a small distance from its equilibrium position E, and is released from rest at time $t = 0$. In the subsequent motion, angle $PCD = \theta$, and ϕ is the angle which PD makes with the tangent at P. Given that θ is small enough for θ^2 and higher powers of θ to be neglected, show that

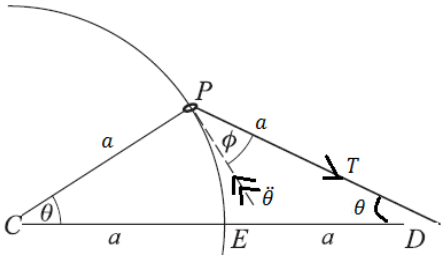
(i) $PD = a$ m, [2]

(ii) the tension T N in the string is constant, [1]

(iii) $\phi = \frac{1}{2}\pi - 2\theta$. [3]

Write down an approximate equation of motion of P on the wire. [2]

The period of the motion of P is approximately 2 s. Find the value of T in terms of m and a, and find the solution of the approximate equation of motion for which $\theta > 0$ when $t = 0$ and, 0.5 s after release, $\dot{\theta} = -0.03$ rad s^{-1}. [6]



(i) Using the law of cosines,

$$PD^2 = a^2 + 4a^2 - 2(2a)a\cos\theta$$

When θ is small, $\cos\theta \approx 1$.

Thus $PD^2 \approx 5a^2 - 4a^2 = a^2$

$$PD \approx a$$

(ii) $PD \approx PE$, force due to extension ≈ 0.
Hence tension is almost constant.

(iii) As $PC = PD, \Delta PCD$ is isosceles.

$$\theta + \theta + \varphi + \frac{\pi}{2} = \pi, \quad \varphi = \frac{1}{2}\pi - 2\theta$$

Resolving tangentially to the circle, $T\cos\varphi = -ma\,\ddot{\theta}$.

As $\cos\varphi = \sin 2\theta \approx 2\theta$, $\ddot{\theta} = -\frac{2T}{ma}\theta$.

The motion is approximately simple harmonic with $\omega^2 = \frac{2T}{ma}$.

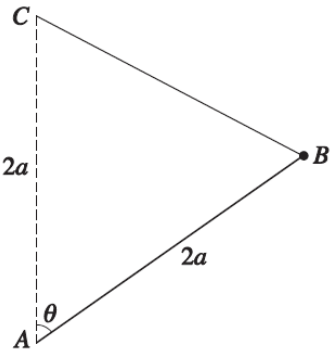
$$\frac{2\pi}{\omega} = 2 \rightarrow \omega^2 = \pi^2, \qquad T = \frac{1}{2}\pi^2 ma$$

Equation of motion is $\theta = A\cos\pi t \rightarrow \dot{\theta} = -\pi A\sin\pi t$.

$$\dot{\theta} = -0.03, t = 0.5: A = \frac{0.03}{\pi}$$

Thus the equation of motion is $\theta = \frac{0.03}{\pi}\cos\pi t$.

Example 6: W2007 P2 Q11 EITHER



A light rod AB of length $2a$ is freely jointed at A to a fixed point and a particle of mass m is attached to B. A light elastic string, of natural length a and modulus of elasticity kmg, has one end attached to B and the other end attached to a point C which is fixed at a distance $2a$ vertically above A (see diagram). The system is in motion in a vertical plane, without the string becoming slack. If air resistance is neglected, show that

$$\ddot{\theta} = \frac{g}{2a}\left(k\cos\tfrac{1}{2}\theta - (2k-1)\sin\theta\right),$$

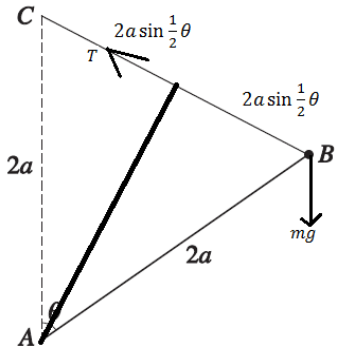
where angle $CAB = \theta$. [6]

In the equilibrium position, $\theta = \frac{1}{3}\pi$. Find the value of k. [2]

By substituting $\theta = \frac{1}{3}\pi + \varepsilon$, where ε is small, show that the motion is approximately simple harmonic, and give the period. [6]

Note: first part of the question requires knowledge from Chapter 14: Rotation of a Rigid Body

Free Body Diagram:



Moment of inertia of the system at A = $m(2a)^2 = 4ma^2$.

$$T = \frac{kmg(BC-a)}{a} = kmg\left(4\sin\frac{1}{2}\theta - 1\right)$$

Using $\tau = I\ddot{\theta}$ at A:

$$4ma^2\ddot{\theta} = mg(2a\sin\theta) - T(2a)\left(\cos\frac{1}{2}\theta\right)$$

$$\frac{2a}{g}\ddot{\theta} = \sin\theta - k\left(4\sin\frac{1}{2}\theta - 1\right)\cos\frac{1}{2}\theta$$

$$= \sin\theta - k\left(2\sin\theta - \cos\tfrac{1}{2}\theta\right)$$

$$\ddot{\theta} = \frac{g}{2a}\left[k\cos\frac{1}{2}\theta - (2k-1)\sin\theta\right]$$

At $\theta = \frac{1}{3}\pi, \ddot{\theta} = 0$:

$$k\cos\frac{1}{6}\pi = (2k-1)\sin\frac{1}{3}\pi \rightarrow k = 2k-1, k = 1$$

$\theta = \frac{1}{3}\pi + \varepsilon, k = 1$:

$$\ddot{\theta} = \frac{g}{2a}\left[\cos\left(\frac{1}{6}\pi + \frac{1}{2}\varepsilon\right) - \sin\left(\frac{1}{3}\pi + \varepsilon\right)\right]$$

$$= \frac{g}{2a}\left(\cos\frac{1}{6}\pi\cos\frac{1}{2}\varepsilon - \sin\frac{1}{6}\pi\sin\frac{1}{2}\varepsilon - \sin\frac{1}{3}\pi\cos\varepsilon - \cos\frac{1}{3}\sin\varepsilon\right)$$

As ε is small, $\sin\varepsilon \rightarrow \varepsilon, \cos\varepsilon \rightarrow 1, \sin\frac{1}{2}\varepsilon \rightarrow \frac{1}{2}\varepsilon, \cos\frac{1}{2}\varepsilon \rightarrow 1$.

$$\ddot{\theta} = \frac{g}{2a}\left[\frac{\sqrt{3}}{2} - \frac{1}{2}\left(\frac{1}{2}\varepsilon\right) - \frac{\sqrt{3}}{2}\varepsilon - \frac{1}{2}\varepsilon\right] \rightarrow \ddot{\theta} = -\frac{3g}{8a}\varepsilon$$

Since $\varepsilon = \theta - \frac{1}{3}\pi, \ddot{\varepsilon} = \ddot{\theta}$ so $\ddot{\varepsilon} = -\frac{3g}{8a}\varepsilon$.

Hence the motion is approximately simple harmonic with $\omega^2 = \frac{3g}{8a}$.

$$T = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{8a}{3g}}$$

CHAPTER 14: ROTATION OF A RIGID BODY

14.1 Moment of Inertia

- LO: understand and use the definition of the moment of inertia of a system of particles about a fixed axis as Σmr^2, and the additive property of moment of inertia for a rigid body composed of several parts (the use of integration to find moments of inertia will not be required);
- LO: use the parallel and perpendicular axes theorems (proofs of these theorems will not be required);

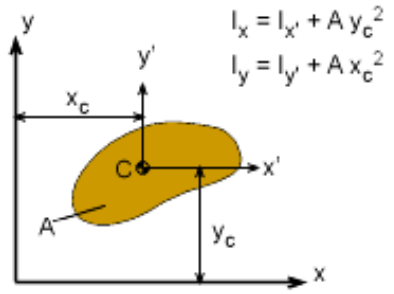
- ❖ Pre-requisite knowledge: Centre of mass (Chapter 12), Circular Motion (Chapter 11), Moments (Chapter 12), all M1 topics regarding linear motion
- ❖ Just like mass is the resistance of an object to a force, moment of inertia is the resistance of an object to a moment/ couple.
- ❖ Moment of inertia, $I = \sum m_i r_i^2$. It's SI unit is kgm^2.
- ❖ Moment of inertia of an object changes depending on where the axis is taken.
- ❖ Moment of inertia of a ring, $I = mr^2$ (treat the ring as multiple point masses and add everything up)
- ❖ Moment of inertia of other objects is given below (quoted from MF10):

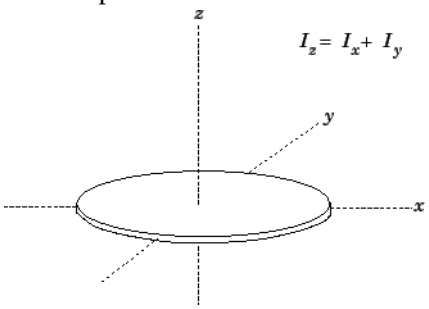
Thin rod, length $2l$, about perpendicular axis through centre: $\frac{1}{3}ml^2$
Rectangular lamina, sides $2a$ and $2b$, about perpendicular axis through centre: $\frac{1}{3}m(a^2+b^2)$
Disc or solid cylinder of radius r about axis: $\frac{1}{2}mr^2$
Solid sphere of radius r about a diameter: $\frac{2}{5}mr^2$
Spherical shell of radius r about a diameter: $\frac{2}{3}mr^2$

Note: Be careful with cases involving rods and laminas, where the distance must be halved!

Properties involving moment of inertia:

- ❖ Can be added / subtracted (when different objects or particles combine, are added or are taken away)
- ❖ Changes according to position where the axis is taken (use parallel / perpendicular axis theorem as quoted below)

Parallel axis theorem	
	➢ If I denotes the moment of inertia of a rigid object of mass M about some axis, and I_G denotes the moment of inertia about the parallel axis through the centre of mass, then $I = I_G + Ma^2$, where a is the perpendicular distance between the two axes. ➢ Used to shift the axis where the moment of inertia is considered from a point to another point.

Perpendicular axis theorem	
	➢ For a rigid object which lies entirely in one plane, the moment of inertia about an axis at right angles to the plane, through a point O of the plane, is equal to sum of moments of inertia about two perpendicular axes in the plane through O. ($I_Z = I_X + I_Y$) ➢ Used to shift the axis of rotation to another axis. ➢ Special case: when the object is symmetrical at the centre, $I_X = I_Y$, hence $I_Z = 2I_X$. ➢ Special case: when object cannot rotate through one of the axes e.g. y-axis, $I_Y = 0$, thus $I_Z = I_X$.

Example 1: S2003 P2 Q5 (first part)

A rigid body of mass M consists of four equal uniform rods joined at their ends to form a square $ABCD$ of side $2a$. Show that the moment of inertia of this body about an axis

(i) through its centre perpendicular to its plane is $\frac{4}{3}Ma^2$, [2]

(ii) through A perpendicular to its plane is $\frac{10}{3}Ma^2$, [2]

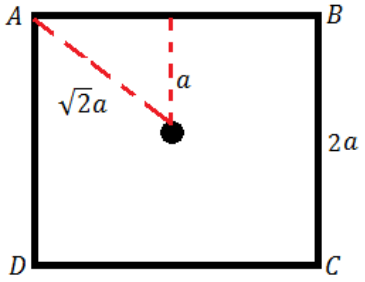
(iii) through A parallel to BD is $\frac{8}{3}Ma^2$. [4]

(i) About the centre, $I_{rod} = \frac{1}{3}\left(\frac{M}{4}\right)(a)^2 + \frac{M}{4}a^2 = \frac{1}{3}Ma^2$

$$I_{system} = 4I_{rod} = \frac{4}{3}Ma^2$$

(ii) About A, use the parallel axis theorem:

$$I_{system} = \frac{4}{3}Ma^2 + M(2a^2) = \frac{10}{3}Ma^2$$



(iii) Perpendicular axis theorem:

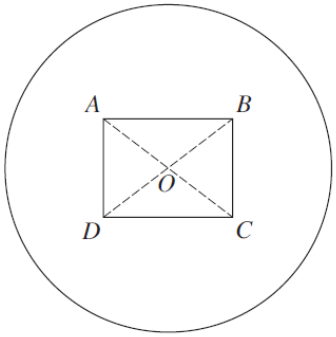
$$2I_X = I_{system},\ I_X = \frac{2}{3}Ma^2$$

About A, use the parallel axis theorem:

$$I_{system} = \frac{2}{3}Ma^2 + M(2a^2) = \frac{8}{3}Ma^2$$

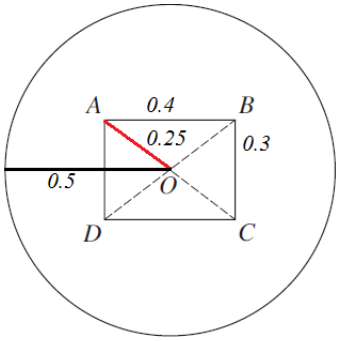
Note: Always use the perpendicular axis theorem if needed before the parallel axis theorem to shift the axis of rotation to another point as symmetry occurs at the centre of the system.

Example 3: S2008 P2 Q5 (first part)



A uniform circular disc has radius 0.5 m, centre O and mass per unit area 6.20 kg m^{-2}. A rectangle $ABCD$, where AC and BD intersect at O, is removed (see diagram). The lengths of AB and BC are 0.4 m and 0.3 m respectively. Show that the moment of inertia of the resulting lamina, about an axis through A perpendicular to its plane, is 0.851 kg m^2, correct to 3 significant figures. [8]

$$\text{Mass of disc} = \pi(0.5)^2(6.20) = 1.55\pi\text{kg}$$
$$\text{Mass of lamina} = 0.4(0.3)(6.20) = 0.744\text{kg}$$



About A, (use parallel axes theorem for both parts)

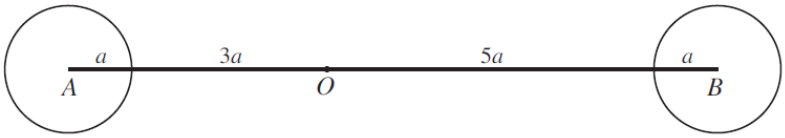
$$I_{disc} = \frac{1}{2}(1.55\pi)(0.5)^2 + 1.55\pi(0.25)^2 = \frac{93}{320}\pi$$

$$\text{I}_{\text{lamina}} = \frac{1}{3}(0.744)(0.2^2 + 0.15^2) + (0.744)(0.25)^2 = \frac{31}{500}$$

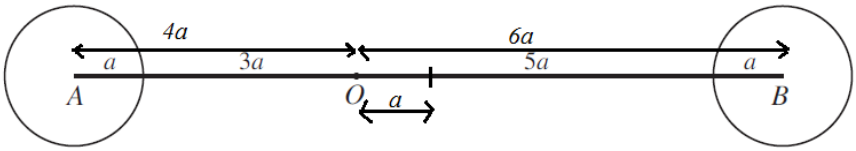
Since the lamina is removed,

$$\text{I}_{\text{system}} = \text{I}_{\text{disc}} - \text{I}_{\text{lamina}} = \frac{93}{320}\pi - \frac{31}{500} \approx 0.851\ \text{kgm}^2$$

Example 3: W2012 P21 Q1



A rigid body consists of two uniform circular discs, each of mass m and radius a, the centres of which are rigidly attached to the ends A and B of a uniform rod of mass $3m$ and length $10a$. The discs and the rod are in the same plane and O is the point on the rod such that $AO = 4a$ (see diagram). Show that the moment of inertia of the body about an axis through O perpendicular to the plane of the discs is $81ma^2$. [5]



About O, (use parallel axes theorem for all three parts)

$$I_{rod} = \frac{1}{3}(3m)(5a)^2 + (3m)a^2 = 28ma^2$$
$$I_A = \frac{1}{2}(m)a^2 + m(4a)^2 = 16\frac{1}{2}ma^2$$
$$I_B = \frac{1}{2}(m)a^2 + m(6a)^2 = 36\frac{1}{2}ma^2$$
$$I_{system} = 28ma^2 + 16\frac{1}{2}ma^2 + 36\frac{1}{2}ma^2 = 81ma^2$$

14.2 Rotational Motion

- LO: recall and use the equation of angular motion $C = I\ddot{\Theta}$ for the motion of a rigid body about a fixed axis (simple cases only, where the moment C arises from constant forces such as weights or the tension in a string wrapped around the circumference of a flywheel; knowledge of couples is not included and problems will not involve consideration or calculation of forces acting at the axis of rotation);
- LO: recall and use the formulae ½ $\text{I}\omega^2$ for the kinetic energy of a rigid body rotating about a fixed axis;
- LO: use considerations of energy in solving problems concerning mechanical systems where motion of a rigid body about a fixed axis is involved;

❖ Conversion between linear motion and rotational motion 'language':

Terms	Linear Motion	Rotational Motion
resistance to motion	mass, m	moment of inertia, $I = \sum m_i r_i^2$
displacement	linear displacement, s	angular displacement, θ
velocity	linear velocity, v	angular velocity, $\omega = \dot{\theta}$
acceleration	linear acceleration, a	angular acceleration, $\alpha = \dot{\omega} = \ddot{\theta}$
Newton's 2nd Law	force, $F = ma$	moment, $C = I\ddot{\theta}$
kinetic energy	$\frac{1}{2}mv^2$	$\frac{1}{2}I\omega^2$
work done	Fs	$C\theta$
work-energy equation:	$Fs = \frac{1}{2}mv^2 + mgh$	$C\theta = \frac{1}{2}I\omega^2 + mg\bar{y}$

❖ Conversion of constant acceleration equations

Linear Motion	Rotational Motion
$v = u + at$	$\omega = \omega_0 + \alpha t$
$s = ut + \frac{1}{2}at^2$	$\theta = \omega_0 t + \frac{1}{2}\alpha t^2$
$s = \frac{1}{2}(u + v)t$	$\theta = \frac{1}{2}(\omega_0 + \omega)t$
$v^2 = u^2 + 2as$	$\omega^2 = {\omega_0}^2 + 2\alpha\theta$
$s = vt - \frac{1}{2}at^2$	$\theta = \omega t - \frac{1}{2}\alpha t^2$

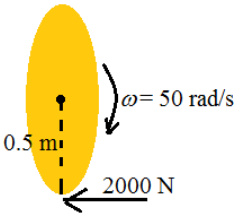
❖ Calculation of gravitational potential energy: The formula $E_p = mgh$ is still used, but distance must be taken from centre of mass of an object.

Problem solving strategies:

1. Find the moment of inertia of the system (remember that it changes and has to be recalculated whenever the axis changes or when extra particles are added/ removed!)
2. To find velocity, take the work-energy equation (remember to always take GPE from distance from centre of mass)
3. Draw free body diagrams to make change in heights for GPE calculation simpler – don't draw the whole diagram, sometimes a line is good enough to represent a lamina!
4. To use Newton's 2nd law / prove SHM: use $C = I\ddot{\theta}$
5. Use constant acceleration equations only when acceleration is confirmed constant
6. For extra connected particles, resolve using $F = ma = mr\ddot{\theta}$
7. Connection to Chapter 13 – prove SHM and find period of oscillation ($\sin\theta \approx \theta$)

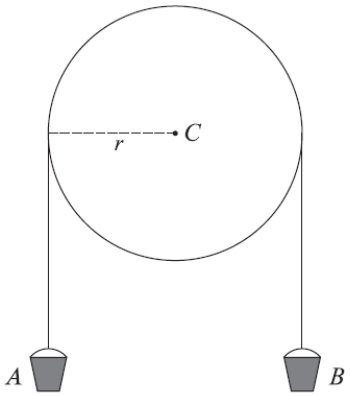
Example 1: S2004 P2 Q2

A circular flywheel of radius 0.5 m, with moment of inertia about its axis of 500 kg m^2, is rotating freely at an angular speed of 50 rad s^{-1}. A brake that exerts a tangential force of constant magnitude 2000 N is applied to its rim. Find the time taken for the flywheel to be brought to rest, and the energy lost in kilojoules. [6]



$$C = I\ddot{\theta} \qquad \ddot{\theta} = \frac{\omega_f - \omega_i}{t}$$
$$-2000(0.5) = 500\ddot{\theta} \qquad t = \frac{-50}{-2} = 25s$$
$$\ddot{\theta} = -2rads^{-2}$$
$$\text{Energy loss} = \frac{1}{2}I\omega^2 = \frac{1}{2}(500)(50)^2 = 625\text{ kJ}$$

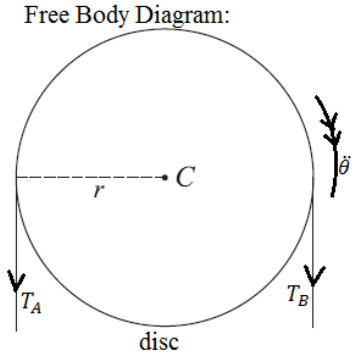
Example 2: W2005 P2 Q3



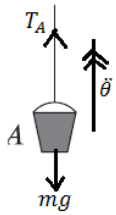
A uniform wheel, in the shape of a circular disc of radius r m, has mass $12m$ kg. It is free to turn in a vertical plane about a smooth fixed horizontal axis through its centre C. A light inextensible rope passes over the upper half of the wheel. Buckets A and B, each of mass m kg, hang freely at the ends of the rope (see diagram). When the wheel is at rest, a weight of mass λm kg is placed gently in B. Find the acceleration with which B begins to move, given that the rope does not slip. [6]

To prevent the rope slipping, the tensions on the two sides of the wheel must not differ by more than mg N. Find the greatest possible value of λ. [3]

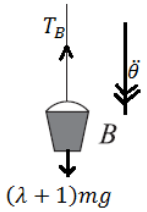
Free Body Diagram:



disc



pail A



pail B

Resolving vertically for B:

$$(\lambda+1)mg - T_B = (\lambda+1)mr\ddot{\theta}$$
$$T_B = (\lambda+1)mg - (\lambda+1)mr\ddot{\theta}$$

Resolving vertically for A:

$$T_A - mg = mr\ddot{\theta}$$
$$T_A = mg + mr\ddot{\theta}$$

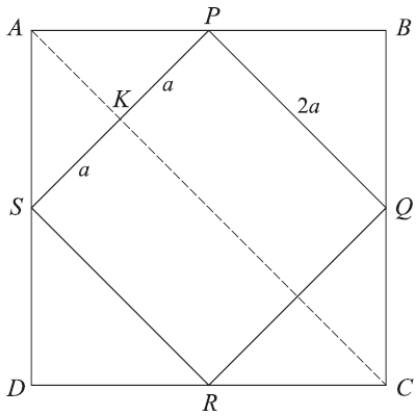
$$I_{disc} = \frac{1}{2}(12m)r^2 = 6mr^2$$

Using $C = I\ddot{\theta}$ for disc:

$$(T_B - T_A)r = I\ddot{\theta}$$
$$\lambda mg - (\lambda+2)mr\ddot{\theta} = 6mr\ddot{\theta}$$
$$mr\ddot{\theta}(8+\lambda) = \lambda mg$$
$$a = r\ddot{\theta} = \frac{\lambda g}{8+\lambda}$$

$$T_B - T_A \le mg$$
$$\lambda mg - \frac{mg\lambda(\lambda+2)}{8+\lambda} \le mg$$
$$8\lambda + \lambda^2 - \lambda^2 - 2\lambda \le 8 + \lambda$$
$$5\lambda \le 8 \Rightarrow \lambda_{max} = \frac{8}{5}$$

Example 3: W2005 P2 Q11 EITHER



Two uniform square laminas are combined into a single body. One lamina $ABCD$ has mass $5m$ and the other lamina $PQRS$ has mass m. The lamina $PQRS$ has side $2a$, and its vertices are at the mid-points of the sides of $ABCD$, with P on AB and S on AD. The line PS meets AC at K, and the body rotates in a vertical plane about a horizontal axis k through K (see diagram). Find the moment of inertia of the body about k. [5]

(i) Assume that there is no resistance to motion. The vertex C is held vertically above K and is slightly displaced so that the body begins to rotate about k, reaching a greatest angular speed ω. Show that $\omega^2 = \frac{9g}{5a}$. [3]

(ii) Assume instead that there is a resistance to motion. The vertex C is held vertically above K and is projected horizontally. When the speed of projection is $3\sqrt{\left(\frac{ga}{5}\right)}$ the body comes to rest, after half a revolution, with C vertically below K. When the speed of projection is v the body comes to rest after making one complete revolution. The work done against the resistance when the body makes a complete revolution is twice the work done when the body makes half a revolution. Find v in terms of g and a. [6]

About O: $I_{ABCD} = \frac{1}{3}(5m)[(2a\cos 45)^2 + (2a\cos 45)^2] = \frac{20}{3}ma^2$

$$I_{PQRS} = \frac{1}{3}m(a^2 + a^2) = \frac{2}{3}ma^2$$

$$I_{\text{system}} = I_{ABCD} + I_{PQRS} = \frac{22}{3}ma^2$$

Using parallel axes theorem, $I_K = \frac{22}{3}ma^2 + 6m(a^2) = \frac{40}{3}ma^2$.

C
a
K
a
C′

(i) Speed of system will reach maximum when C is vertically below K.

Using conservation of energy: $\frac{1}{2}I\omega^2 = mgh$

$$\frac{1}{2}\left(\frac{40}{3}ma^2\right)\omega^2 = 6mg(2a)$$

$$\frac{20}{3}a\omega^2 = 12g \rightarrow \omega^2 = \frac{9g}{5a}(SHOWN)$$

(ii) Let the work done be W. For half a revolution, $r = 3a$, $\omega = \frac{v}{r} = \sqrt{\frac{g}{5a}}$.

For half a revolution: $W = loss\ in\ E_k + loss\ in\ E_p$

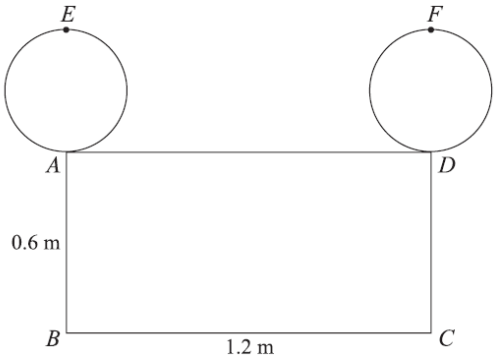
$$= \frac{1}{2}\left(\frac{40}{3}ma^2\right)\left(\frac{g}{5a}\right) + 6mg(2a) = \frac{40}{3}mga$$

For a full revolution: $2W = loss\ in\ E_k\ (E_p loss = 0)$

$$\frac{80}{3}mga = \frac{1}{2}\left(\frac{40}{3}ma^2\right)\omega^2 \rightarrow \omega^2 = \frac{4g}{a}$$

$$v = r\omega = 3a\sqrt{\frac{4g}{a}} = 6\sqrt{ga}$$

Example 4: W2006 P2 Q11 EITHER



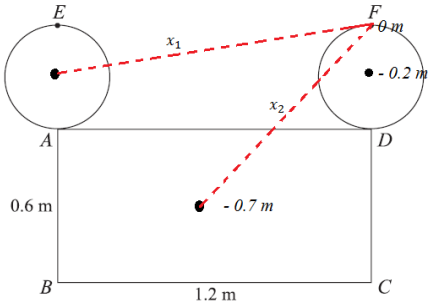
A sign is formed from two uniform discs, each of mass 0.25 kg and radius 0.2 m, rigidly fixed to a uniform rectangular lamina $ABCD$ at A and D. The disc attached at A has diameter AE and BAE is a straight line. The disc attached at D has diameter DF and CDF is a straight line. The discs are in the same plane as $ABCD$. The mass of the rectangular lamina is M kg, and $AB = 0.6$ m, $BC = 1.2$ m. The sign is hung from smooth supports at E and F with AD horizontal (see diagram). Show that the moment of inertia of the sign about an axis through F perpendicular to $ABCD$ is $(0.39 + M)$ kg m^2. [8]

The support at E breaks and the sign rotates freely under gravity, in a vertical plane, about the axis through F. Assuming that AD becomes vertical in the subsequent motion when the sign is rotating with angular speed ω rad s^{-1}, show that

$$\omega^2 = \frac{2(2-M)}{0.39+M}.$$ [4]

What happens if $M > 2$? [2]

Initial:



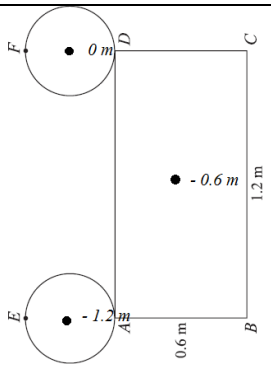
Applying Pythagoras' theorem:

$$x_1^2 = (0.2)^2 + (1.2)^2 = 1.48\ m^2$$
$$x_2^2 = (0.7)^2 + (0.6)^2 = 0.85\ m^2$$

About F (use parallel axes theorem):

$$\mathrm{I_{ABCD}} = \frac{1}{3}M(0.3^2 + 0.6^2) + Mx_2^2 = M\ \mathrm{kgm^2}$$
$$\mathrm{I_{DF}} = (0.5)(0.25)(0.2)^2 + (0.25)x_3^2 = 0.015\ \mathrm{kgm^2}$$
$$\mathrm{I_{AE}} = (0.5)(0.25)(0.2)^2 + (0.25)x_1^2 = 0.375\ \mathrm{kgm^2}$$
$$\mathrm{I_{system}} = \mathrm{I_{AE}} + \mathrm{I_{DF}} + \mathrm{I_{ABCD}} = (0.39 + M)\mathrm{kgm^2}$$

Final:



Setting F as 0 m;

Δh for disc DF $= 0 - (-0.2) = 0.2\ m$

Δh for disc EA $= -1.2 - (-0.2) = -1.0\ m$

Δh for lamina ABCD $= -0.6 - (-0.7) = 0.1\ m$

Conservation of energy: Gain in KE = Loss in GPE

$$\frac{1}{2}(0.39 + M)\omega^2$$
$$= 0.25(10)(1) - 10M(0.1) - 10(0.25)(0.2)$$
$$\frac{1}{2}(0.39 + M)\omega^2 = 2 - M \rightarrow \omega^2 = \frac{2(2-M)}{0.39+M}$$

If $M > 2$, ω^2 is negative when AD is vertical.

Since the sign stops rotating upwards when $\omega = 0$, it will not rotate until AD is vertical.

Example 5: W2009 P2 Q4

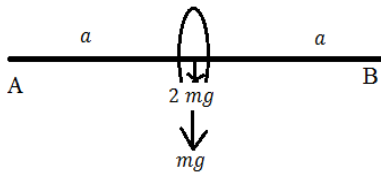
A uniform rod AB, of length $2a$ and mass $2m$, can rotate freely in a vertical plane about a smooth horizontal axis through A. A small rough ring of mass m is threaded on the rod. The rod is held in a horizontal position with the ring at rest at the mid-point of the rod. The rod is released from rest. Using energy considerations, show that, until the ring slides,

$$a\dot{\theta}^2 = \tfrac{18}{11}g\sin\theta,$$

where θ is the angle turned through by the rod. [3]

Show that, until the ring slides, the magnitudes of the friction force and normal contact force acting on the ring are $\frac{29}{11}mg\sin\theta$ and $\frac{2}{11}mg\cos\theta$ respectively. [6]

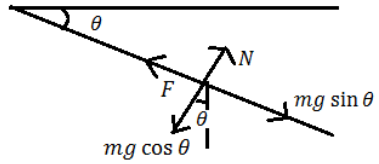
The coefficient of friction between the ring and the rod is μ. Find, in terms of μ, the value of θ when the ring starts to slide. [2]



$$I = \frac{1}{3}(2m)a^2 + ma^2 + 2ma^2 = \frac{11}{3}ma^2$$

$$\frac{1}{2}I\dot{\theta}^2 = mgh \rightarrow \frac{1}{2}\left(\frac{11}{3}ma^2\right)\dot{\theta}^2 = 3mg(a\sin\theta)$$

$$ma^2\dot{\theta}^2 = \frac{18}{11}mga\sin\theta \rightarrow a\dot{\theta}^2 = \frac{18}{11}g\sin\theta \text{ (SHOWN)}$$



Resolve along the rod: $F - mg\sin\theta = ma\dot{\theta}^2$

$$F = \frac{18}{11}mg\sin\theta + mg\sin\theta = \frac{29}{11}mg\sin\theta \text{ (SHOWN)}$$

Differentiating with respect to time (use the chain rule on both sides)

$$a\dot{\theta}^2 = \frac{18}{11}g\sin\theta \rightarrow 2a\dot{\theta}\ddot{\theta} = \frac{18}{11}g\cos\theta\,\dot{\theta} \rightarrow \ddot{\theta} = \frac{9g}{11a}\cos\theta$$

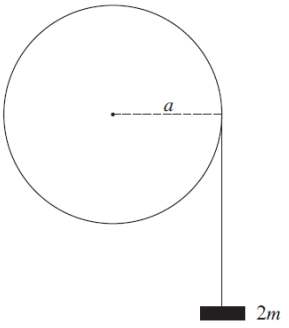
Resolve perpendicularly to the rod, $mg\cos\theta - N = ma\ddot{\theta}$

$$N = mg\cos\theta - \frac{9}{11}mg\cos\theta = \frac{2}{11}mg\cos\theta\ (SHOWN)$$

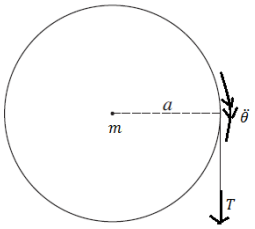
When ring starts to slide, $F = \mu N$.

$$\frac{29}{11}mg\sin\theta = \frac{2}{11}mg\cos\theta\,(\mu) \rightarrow \mu = \frac{29}{2}\tan\theta \rightarrow \theta = \tan^{-1}\left(\frac{2}{29}\mu\right)$$

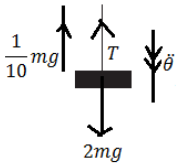
Example 6: W2008 P2 Q3



A uniform disc, of mass m and radius a, is free to rotate without resistance in a vertical plane about a horizontal axis through its centre. A light inextensible string has one end fixed to the rim of the disc, and is wrapped round the rim. A block of mass $2m$ is attached to the other end of the string (see diagram). The system is released from rest with the block hanging vertically. While the block moves it experiences a constant resistance to motion of magnitude $\frac{1}{10}mg$. Find the angular acceleration of the disc, and find also the angular speed of the disc when it has turned through one complete revolution. [9]



Using $C = I\ddot{\theta}$ for disc: $Ta = \frac{1}{2}ma^2\ddot{\theta} - (1)$



Using $F = ma$ for block:

$$2mg - T - \frac{1}{10}mg = 2ma\ddot{\theta} - (2)$$

---

(1) into (2): $\frac{19}{10}mg - \frac{1}{2}ma\ddot{\theta} = 2ma\ddot{\theta}$

$$\frac{5}{2}ma\ddot{\theta} = \frac{19}{10}mg \rightarrow \ddot{\theta} = \frac{19g}{25a}$$

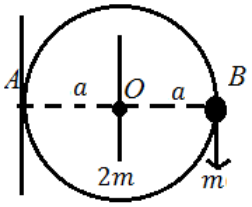
Using $\omega^2 = {\omega_0}^2 + 2\alpha\theta, \omega^2 = 2\left(\frac{19g}{25a}\right)(2\pi) = \frac{76g\pi}{25a} \rightarrow \omega = \frac{9.77}{\sqrt{a}}$ (3 s. f.)

Example 7: W2010 P2 Q5

A uniform circular disc has diameter AB, mass $2m$ and radius a. A particle of mass m is attached to the disc at B. The disc is able to rotate about a smooth fixed horizontal axis through A. The axis is tangential to the disc. Show that the moment of inertia of the system about the axis is $\frac{13}{2}ma^2$. [4]

The disc is held with AB horizontal and released. Find the angular speed of the system when B is directly below A. [5]

The disc is slightly displaced from the position of equilibrium in which B is below A. At time t the angle between AB and the vertical is θ. Write down the equation of motion, and find the approximate period of small oscillations about the equilibrium position. [5]



Using perpendicular axis theorem at O, ($I_X = I_Y,\ I_Y = 2I_Z$)

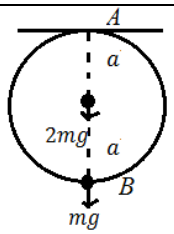
$$I_{disc} = \frac{1}{2}\left(\frac{1}{2}\right)(2m)a^2 = \frac{1}{2}ma^2$$

About A, (parallel axis theorem for disc)

$$I_{disc} = \frac{1}{2}ma^2 + 2m(a^2) = \frac{5}{2}ma^2$$

$$I_{particle} = m(2a)^2 = 4ma^2$$

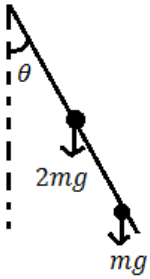
$$I_{system} = I_{disc} + I_{particle} = \frac{13}{2}ma^2$$



Using conservation of energy: Gain in KE = Loss in GPE

$$\frac{1}{2}\left(\frac{13}{2}ma^2\right)\omega^2 = 2mga + 2mga$$

$$\frac{13}{4}a\omega^2 = 4g \rightarrow \omega = \sqrt{\frac{16g}{13a}}$$



Using $C = I\ddot{\theta}$:

$$\frac{13}{2}ma^2\ddot{\theta} = 2mga\sin\theta + 2mga\sin\theta$$

$$\ddot{\theta} = \frac{8g}{13a}\sin\theta$$

When θ is small, $\sin\theta \approx \theta$, thus $\ddot{\theta} \approx \frac{8g}{13a}\theta$.

The motion is approximately simple harmonic with $\omega^2 = \frac{8g}{13a}$.

$$T = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{13a}{8g}}$$

FURTHER MATHEMATICS PAPER 2 STATISTICS

CHAPTER 15: FURTHER WORK ON DISTRIBUTIONS

15.1 Pre-Requisite Knowledge:

A) Probability Distributions (P6, P7)

1. Identities for Discrete Probability Distributions (P6):

$$\sum_{\text{all } x} P(X = x) = 1, \qquad E(X) = \sum_{\text{all } x} xP(X = x),$$

$$E[g(X)] = \sum_{\text{all } x} g(x)P(X = x), \qquad Var(X) = E(X^2) - [E(X)]^2$$

2. Special Discrete Probability Distributions (P6, P7)

Name	Parameter	Formula	E(X)	Var(X)
Binomial Distribution	X~ B(n, p)	$P(X = r) = {}^nC_r p^r q^{n-r}, q = 1 - p$	np	npq
Poisson Distribution	X~ Po(λ)	$P(X = r) = e^{-\lambda}\frac{\lambda^r}{r!}$	λ	λ

3. Normal Distribution (P6)

Normal Variable: $X \sim N(\mu, \sigma^2)$

Standardised variable: $Z \sim N(0,1)$. To standardise, use $Z = \frac{X - \mu}{\sigma}$

4. Approximations

Approximations	To	How	Condition
Normal Approximation	Binomial	$X \sim B(n, p) \rightarrow X \sim N(np, npq)$	$np > 5, nq > 5$
Poisson Approximation	Binomial	$X \sim B(n, p) \rightarrow X \sim Po(np)$	$n > 50$, $np < 5$
Normal Approximation	Poisson	$X \sim Po(\lambda) \rightarrow X \sim N(\lambda, \lambda)$	$\lambda > 15$

5. Others

Basic Probability Concepts (P6 Conditional Probability, etc.)

Continuity Correction: Needed when changing from discrete to continuous. Examples:

$P(X < 32) \rightarrow P(X < 31.5)$	$P(X \leq 32) \rightarrow P(X < 32.5)$
$P(X > 32) \rightarrow P(X > 32.5)$	$P(X \geq 32) \rightarrow P(X > 31.5)$
$P(3 < X < 10) \rightarrow P(3.5 < X < 9.5)$	$P(3 \leq X \leq 10) \rightarrow P(2.5 < X < 10.5)$

B) Linear Combinations of Random Variables (P7):

1. For Two Independent Discrete Random Variables:

$$E(aX) = aE(X) \qquad Var(aX) = a^2Var(X)$$

$$E(aX + bY) = aE(X) + bE(Y) \qquad Var(aX + bY) + a^2Var(X) + b^2Var(Y)$$

$$E(aX - bY) = aE(X) - bE(Y) \qquad Var(aX - bY) + a^2Var(X) + b^2Var(Y)$$

$$E(X_1 + X_2 + ... + X_n) = nE(X) \qquad Var(X_1 + X_2 + ... + X_n) = nVar(X)$$

Note: The difference between repeated observations and multiples of observations

Note: Variance is always added!

2. Special Probability Distribution Model Combinations:

If $X \sim Po(\lambda_1), Y \sim Po(\lambda_2)$, then $X + Y \sim Po(\lambda_1 + \lambda_2)$

If $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$, then

$$X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2) \qquad aX \sim N\left(a\mu, a^2\sigma^2\right)$$

$$X - Y \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2) \qquad aX + b \sim N(a\mu + b, a^2\sigma^2)$$

$$X_1 + X_2 + ... + X_n \sim N(n\mu, n\sigma^2) \qquad aX + bY \sim N(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)$$

$$aX - bY \sim N(a\mu_1 - b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)$$

Example 1: The Poisson distribution + Approximation (S2003 Mathematics P7 Q6)

Computer breakdowns occur randomly on average once every 48 hours of use.

(i) Calculate the probability that there will be fewer than 4 breakdowns in 60 hours of use. [3]

(ii) Find the probability that the number of breakdowns in one year (8760 hours) of use is more than 200. [4]

(iii) Independently of the computer breaking down, the computer operator receives phone calls randomly on average twice in every 24-hour period. Find the probability that the total number of phone calls and computer breakdowns in a 60-hour period is exactly 4. [3]

(i) Let X_t be the number of computer breakdowns in t hours. $X_t \sim Po\left(\frac{1}{48}t\right)$.

$X_{60} \sim Po\,(1.25)$

$$P(X_{60} < 4) = P(X_{60} = 0) + P(X_{60} = 1) + P(X_{60} = 2) + P(X_{60} = 3)$$

$$= e^{-1.25}\left(\frac{1.25^0}{0!} + \frac{1.25^1}{1!} + \frac{1.25^2}{2!} + \frac{1.25^3}{3!}\right)$$

$$\approx 0.962$$

(ii) $X_{8760} \sim Po\,(182.5)$. Using a normal approximation,

$$X_{8760} \sim Po\,(182.5) \rightarrow Y \sim N\,(182.5, 182.5)$$

$$P(X_{8760} > 200) \to P(Y > 200.5)$$
$$= P\left(Z > \frac{200.5 - 182.5}{\sqrt{182.5}}\right)$$
$$= P(Z > 1.332) \approx 0.0915$$

(iii) Let Y_t be the number of phone calls received in t hours. $Y_t \sim Po\left(\frac{2}{24}t\right)$.

$$X_{60} \sim Po\,(1.25), \qquad Y_{60} \sim Po\,(5) \to \; X_{60} + Y_{60} \sim Po\,(6.25)$$
$$P(X_{60} + Y_{60} = 4) = e^{-6.25}\left(\frac{6.25^4}{4!}\right) \approx 0.123$$

Example 2: S2010 Mathematics P73 Q6

Yu Ming travels to work and returns home once each day. The times, in minutes, that he takes to travel to work and to return home are represented by the independent random variables W and H with distributions $N(22.4, 4.8^2)$ and $N(20.3, 5.2^2)$ respectively.

(i) Find the probability that Yu Ming's total travelling time during a 5-day period is greater than 180 minutes. [4]

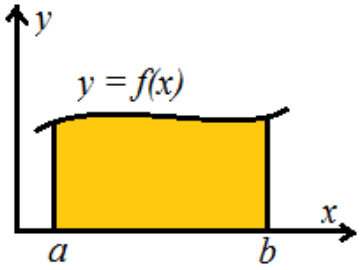
(ii) Find the probability that, on a particular day, Yu Ming takes longer to return home than he takes to travel to work. [5]

(i)
$$5(W + H) \sim N[5(22.4 + 20.3), 5(4.8^2 + 5.2^2)]$$
$$= (213.5, 250.4)$$
$$P[5(W + H) > 180] = P\left(Z > \frac{180 - 213.5}{\sqrt{250.4}}\right)$$
$$= P(Z > -2.117)$$
$$= 0.983$$

(ii)
$$W - H \sim N(22.4 - 20.3, 4.8^2 + 5.2^2) = (2.1, 50.08)$$
$$P(W - H < 0) = P\left(Z < \frac{-2.1}{\sqrt{50.08}}\right)$$
$$= P(Z < -0.297)$$
$$= 0.383$$

15.2 Continuous Random Variables

A) Pre-requisite Knowledge: Probability Density Functions (PDF) (P7)



- Continuous random variables are described using PDFs.
  $y = f(x),\ \ a \le x \le b$

Properties:

1. $\int_a^b f(x)\,dx = 1$
2. For $a \le x_1 \le x_2 \le b, P(x_1 \le X \le x_2) = \int_{x_1}^{x_2} f(x)\,dx$
3. $E(X) = \int_{-\infty}^{\infty} x\,f(x)\,dx$
4. $Var(X) = \int_{-\infty}^{\infty} x^2\,f(x)\,dx - [E(X)]^2$
5. Mode, m: values that satisfy $f'(m) = 0\ , f''(m) < 0$
6. Median, M: $\int_a^M f(x)dx = 0.5$
7. n^{th}percentile, p: $\int_a^p f(x)dx = \frac{n}{100}$

Example 1: Continuous Random Variables (W2004 Mathematics P7 Q6)

A continuous random variable X has probability density function given by

$$f(x) = \begin{cases} 3(1-x)^2 & 0 \leqslant x \leqslant 1, \\ 0 & \text{otherwise.} \end{cases}$$

Find

(i) $P(X > 0.5)$, [3]

(ii) the mean and variance of X. [6]

(i)
$$P(X > 0.5) = \int_{0.5}^{1} 3(1-x)^2 = \left[-(1-x)^3\right]_{0.5}^{1} = (1-0.5)^3 = \frac{1}{8}$$

(ii)
$$E(X) = \int_0^1 3x(1-x)^2 dx = \int_0^1 3x - 6x^2 + 3x^3 dx = \left[\frac{3}{2}x^2 - 2x^3 + \frac{3}{4}x^4\right]_0^1 = \frac{3}{2} - 2 + \frac{3}{4} = \frac{1}{4}$$

$$E\left(X^2\right) = \int_0^1 3x^2(1-x)^2 dx = \int_0^1 3x^2 - 6x^3 + 3x^4 dx = \left[x^3 - \frac{3}{2}x^4 + \frac{3}{5}x^5\right]_0^1 = 1 - \frac{3}{2} + \frac{3}{5} = \frac{1}{10}$$

$$Var(X) = \frac{1}{10} - \left(\frac{1}{4}\right)^2 = \frac{3}{80}$$

Example 2: Continuous Random Variables (S2008 Mathematics P7 Q7)

If Usha is stung by a bee she always develops an allergic reaction. The time taken in minutes for Usha to develop the reaction can be modelled using the probability density function given by

$$f(t) = \begin{cases} \frac{k}{t+1} & 0 \leqslant t \leqslant 4, \\ 0 & \text{otherwise,} \end{cases}$$

where k is a constant.

(i) Show that $k = \frac{1}{\ln 5}$. [4]

(ii) Find the probability that it takes more than 3 minutes for Usha to develop a reaction. [3]

(iii) Find the median time for Usha to develop a reaction. [3]

(i)

$$\int_0^4 \frac{k}{t+1} = 1$$
$$k\left[\ln(t+1)\right]_0^4 = 1$$
$$k(\ln 5 - \ln 1) = 1$$
$$k = \frac{1}{\ln 5}$$

(ii)

$$P(T > 3)$$
$$= \int_3^4 \frac{k}{t+1}$$
$$= k\left[\ln(t+1)\right]_3^4$$
$$= \frac{1}{\ln 5}(\ln 5 - \ln 4)$$
$$\approx 0.139$$

(iii)

$$\int_0^m \frac{k}{t+1} = 0.5$$
$$\left[\ln(t+1)\right]_0^m = 0.5\ln 5$$
$$\ln(m+1) = \ln\sqrt{5}$$
$$m = \sqrt{5} - 1$$
$$\approx 1.24$$

B) Cumulative Distribution Functions / Distribution Functions (CDF)

PROBABILITY DENSITY FUNCTION, f(x) — INTEGRATE → / ← DIFFERENTIATE — CUMULATIVE DISTRIBUTION FUNCTION, F(x)

Properties of the cumulative distribution function:

1. $F(t) = P(X < t) = \int_{-\infty}^{t} f(x)\,dx$
2. $f(x) = \frac{d}{dx}[F(x)]$
3. For $a \le x_1 \le x_2 \le b$,

$$F(a) = 0, \qquad F(b) = 1$$

$$P(x_1 \le X \le x_2) = \int_{x_1}^{x_2} f(x)\,dx = \int_{-\infty}^{x_2} f(x)\,dx - \int_{-\infty}^{x_1} f(x)\,dx = F(x_2) - F(x_1)$$

4. Median, m: $\int_a^M f(x)dx = 0.5 \rightarrow F(m) = 0.5$
5. n^{th}percentile, p: $\int_a^p f(x)dx = \frac{n}{100} \rightarrow F(p) = \frac{n}{100}$

For piece-wise defined functions:

Expectation, variance, probability and distribution functions are still calculated using the PDF equations, but now the integration process must be done by several parts for each different function in the interval.

C) Transformation of Continuous Random Variables

- LO: use the definition of the distribution function as a probability to deduce the form of a distribution function in simple cases, e.g. to find the distribution function for *Y*, where Y = X^3 and *X* has a given distribution;

Technique to solve typical exam questions:

Problem: Given the PDF of X, $f(x)$, find the PDF of Y, $f(y)$, where Y is a given function of X.

Step 1: Find the CDF of *X* (a dummy variable must be used for the integration)

Step 2: Work out the intervals/ limits of the new PDF using the limits of the old PDF.

Step 3: Find the CDF of Y. $F(Y) = P(Y < y)$ then substitute the given function to replace *Y*.

Step 4: Rearrange the formula given to turn the formula into $F(Y) = P\big(X < g(y)\big) = F[g(y)]$

Step 5: Substitute $F(X)$ into the CDF, replacing x with $g(y)$ to obtain $F(Y)$.

Step 6: Differentiate $F(Y)$ to obtain the PDF of Y.

- For piecewise-defined PDF, apply the following steps to each interval given.
- The following steps might appear vague so learn how to use them in the examples given.
- Extra Contents: In general, if $X = f(x)$ and $Y = g(X)$, $E(g(X)) = E(Y)\int_{-\infty}^{\infty} g(x)\, f(x)\, dx$

  Most notably, $E(X^2) = \int_{-\infty}^{\infty} x^2\, f(x)\, dx$.

Example 2: S2007 P2 Q10

The continuous random variable X has probability density function given by

$$f(x) = \begin{cases} 0 & x < 3, \\ 2e^{-(2x-6)} & x \geqslant 3. \end{cases}$$

Obtain the distribution function of Y, where $Y = e^X$. [6]

Find the probability density function of Y and hence find the value of $E(\sqrt{Y})$. [5]

Step 1: CDF of x $F(T) = \int_3^t 2e^{6-2x}\, dx$ $= -\frac{1}{2}[2e^{6-2x}]_3^t$ $= 1 - e^{6-2t}$ $F(x) = 1 - e^{6-2x}, \quad x \geq 3$	Step 4: Substitute $P(Y < y) = P(e^X < y)$ $= P(X < \ln y)$ $= F(\ln y)$
Step 2: Limits $x = 3, y = e^3$	Step 5: Form new CDF $F(\ln y) = 1 - e^{6-2\ln y}$ $F(Y) = \begin{cases} 0, & y \leq e^3 \\ 1 - \frac{e^6}{y^2}, & y \geq e^3 \end{cases}$
Step 3: CDF of y $F(Y) = P(Y < y)$	Step 6: Differentiate for PDF $f(y) = \frac{d}{dy}[1 - e^6 y^{-2}] = 2e^6 y^{-3}$ $f(y) = \begin{cases} \frac{2e^6}{y^3}, & y \geq e^3 \\ 0, & \text{otherwise} \end{cases}$

$$E\left(\sqrt{Y}\right) = \int_{\text{all } y} \sqrt{y}\, f(y) dx$$

$$= \int_{e^3}^{\infty} 2e^6 y^{-\frac{5}{2}} dy$$

$$= -\frac{4}{3} e^6 \left[y^{-1.5}\right]_{e^3}^{\infty}$$

$$= -\frac{4}{3} e^6 \left(0 - e^{-4.5}\right) = \frac{4}{3} e^{\frac{3}{2}}$$

Example 2: S2006 P2 Q6

A car makes a regular journey of 40 km, but owing to varying traffic the time for the journey varies. The time, T hours, may be modelled by a random variable with probability density function given by

$$f(t) = \begin{cases} \frac{1}{2t^3} & t \geqslant \frac{1}{2}, \\ 0 & \text{otherwise.} \end{cases}$$

The average speed for the journey is A km h^{-1}.

(i) Find the distribution function of T and show that the distribution function of A is given, for $0 \leqslant a \leqslant 80$, by

$$G(a) = \frac{a^2}{6400}.$$ [5]

(ii) Find the probability that the average speed for a randomly chosen journey exceeds 60 km h^{-1}. [2]

(i)

$$F(X) = \int_{0.5}^{x} 0.5t^{-3}dt$$
$$= \left[-0.25t^{-2}\right]_{0.5}^{x}$$
$$= 1 - 0.25x^{-2}$$
$$F(T) = 1 - 0.25t^{-2}$$

$$A = \frac{40}{T}$$
$$P(A \leq a) = P\left(\frac{40}{T} \leq a\right)$$
$$F(A) = P\left(T > \frac{40}{a}\right)$$
$$= 1 - F\left(\frac{40}{a}\right)$$
$$= 1 - \left[1 - 0.25\left(\frac{40}{a}\right)^{-2}\right]$$
$$= \frac{0.25a^2}{1600} = \frac{a^2}{6400}, 0 < a \leq 80$$

(ii)

$$1 - G(60)$$
$$= 1 - \frac{3600}{6400}$$
$$= \frac{7}{16}$$

Example 3: S2010 P21 Q11 OR

The continuous random variable T has probability density function given by

$$f(t) = \begin{cases} 0 & t < 2, \\ \frac{2}{(t-1)^3} & t \geqslant 2. \end{cases}$$

(i) Find the distribution function of T, and find also $P(T > 5)$. [3]

(ii) Consecutive independent observations of T are made until the first observation that exceeds 5 is obtained. The random variable N is the total number of observations that have been made up to and including the observation exceeding 5. Find $P(N > E(N))$. [3]

(iii) Find the probability density function of Y, where $Y = \frac{1}{T-1}$. [8]

(i)

$$F(X) = \int_{2}^{x} 2(t-1)^{-3}dt$$
$$= \left[-(t-1)^{-2}\right]_{2}^{x}$$
$$= 1 - \frac{1}{(x-1)^2}$$
$$F(T) = 1 - \frac{1}{(t-1)^2}$$

$$P(T > 5)$$
$$= 1 - F(5)$$
$$= \frac{1}{4^2}$$
$$= \frac{1}{16}$$

(ii) * Involves 15.3 Geometric Distribution

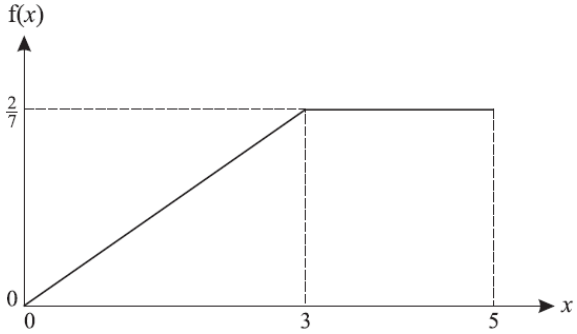
$$N \sim Geo\left(\frac{1}{16}\right)$$
$$E(N) = \frac{1}{\frac{1}{16}} = 16$$
$$P(N > E(N)) = P(N > 16)$$
$$= \left(\frac{15}{16}\right)^{16} = 0.356$$

(iii)

$$P(Y<y)=P\left(\frac{1}{T-1}<y\right)$$
$$=P(T>\frac{1}{y}+1)$$
$$=1-F\left(\frac{1}{y}+1\right)$$
$$=\left(\frac{1}{y}+1-1\right)^{-2}=y^2, 0\le y\le 1$$

$$f(y)=\frac{d}{dy}\left(y^2\right)=2y$$
$$f(y)=\begin{cases}2y, & 0\le y\le 1\\ 0, & \text{otherwise}\end{cases}$$

Example 4: W2006 P2 Q11 OR



The continuous random variable X takes values in the interval $0 \leqslant x \leqslant 5$ only. On this interval the graph of its probability density function, f, consists of the two straight line segments shown in the diagram. Find the distribution function of X. [4]

The random variable Y is given by $Y = \sqrt{(5 - X)}$. Find the distribution function of Y. [6]

Find $P(Y < 2X)$. [4]

For $0 \le x \le 3$:

$$f(x)=\frac{2}{7(3)}x=\frac{2}{21}x$$
$$F(T)=\int_0^t \frac{2}{21}x\,dx=\frac{1}{21}t^2$$

For $3 \le x \le 5$: $f(x)=\frac{2}{7}$

$$F(T)=\int_3^t \frac{2}{7}dx-F(3)$$
$$=\frac{2}{7}t-\frac{3}{7}$$

$$F(x)=\begin{cases}0, & x\le 0\\ \frac{1}{21}x^2, & 0\le x\le 3\\ \frac{2x-3}{7}, & 3\le x\le 5\\ 1, & x\ge 5\end{cases}$$

$$x\le 0, y\ge\sqrt{5}$$
$$0\le x\le 3, \sqrt{2}\le y\le\sqrt{5}$$
$$3\le x\le 5, 0\le y\le\sqrt{2}$$

$$P(Y<y)=P(5-X<y^2)$$
$$=P(X>5-y^2)$$
$$=1-F(5-y^2)$$
$$=\begin{cases}1-\frac{1}{21}(5-y^2)^2, 0\le y\le\sqrt{2}\\ 1-\frac{2(5-y^2)-3}{7}, \sqrt{2}\le y\le\sqrt{5}\\ 1, \quad y\ge\sqrt{5}\end{cases}$$
$$=\begin{cases}\frac{10y^2-y^4-4}{21}, 0\le y\le\sqrt{2}\\ \frac{2}{7}y^2, \quad \sqrt{2}\le y\le\sqrt{5}\\ 1, \quad y\ge\sqrt{5}\end{cases}$$

$$Y<2X$$
$$5-X<4X^2$$
$$(4X+5)(X-1)<0$$
$$X<-\frac{5}{4}, X>1$$

$$P(Y<2X)=P(X>1)$$
$$=1-\frac{1}{21}$$
$$=\frac{20}{21}$$

15.3 Geometric Distribution

- LO: understand conditions under which a geometric distribution may be a suitable probability model;
- LO: recall and use the formula for the calculation of geometric probabilities;
- LO: recall and use the means and variances of a geometric distribution.

- ❖ The geometric distribution is a discrete probability distribution with parameter $X \sim$ Geo (p)
- ❖ Formula: $P(X = r) = (1-p)^{r-1} p$
- ❖ r is no. of trials required for the first successful outcome, p is probability of success
- ❖ Conditions to use it is the same as the binomial distribution: events must be independent and probability of success is constant for each event)
- ❖ $E(X) = \frac{1}{p}, Var(X) = \frac{1-p}{p^2}$
- ❖ Useful Results: $P(X > r) = (1-p)^r$, $P(X \le r) = 1-(1-p)^r$
- ❖ Sometimes involves 2 players (do case-by case analysis / compare expected values)

Example 1: W2003 P2 Q9

Alec and Bill take alternate turns at kicking a football at a goal, and their probabilities of scoring a goal on each kick are p_1 and p_2 respectively, independently of previous outcomes. The first person to score allows the other person one more kick. If the other then scores, the game is drawn. If the other then misses, the first has won the game. Alec begins a game. Show that the probability that

(i) Alec scores first, on his nth kick, is $p_1(q_1q_2)^{n-1}$, where $q_1 = 1 - p_1$ and $q_2 = 1 - p_2$, [2]

(ii) Alec wins the game is $\frac{p_1q_2}{1-q_1q_2}$, [3]

(iii) the game is drawn is $\frac{p_1p_2(1+q_1)}{1-q_1q_2}$. [3]

Given that $p_1 = p_2 = \frac{1}{3}$, find the expected total number of kicks in the game. [2]

(i) $P(\text{Alec scores in nth kick}) = P(\text{Alec miss n -1 times, Bill miss n -1 times, Alec scores})$

$$= q_1^{n-1}q_2^{n-1}p_1 = p_1(q_1q_2)^{n-1}$$

(ii) $P(\text{Alec wins}) = P(AB'+A'B'AB'+A'B'A'B'AB'+...)$

$$= p_1q_2 + p_1q_1q_2^2 + p_1q_2^2q_2^3 + ... = \frac{p_1q_2}{1-q_1q_2}$$

(iii) $P(Draw) = P(AB + A'BA + A'B'AB + A'B'A'BA + ...)$

$$= p_1p_2 + p_1p_2q_1 + p_1p_2q_1q_2 + p_1p_2q_1^2q_2 + ...$$
$$= .p_1p_2(1 + q_1 + q_1q_2 + q_1^2q_2)$$
$$= p_1p_2(1 + q_1q_2 + ...) + p_1p_2(q_1 + q_1^2q_2 + ...)$$
$$= p_1p_2\left(\frac{1}{1-q_1q_2}\right) + p_1p_2\left(\frac{q_1}{1-q_1q_2}\right)$$
$$= p_1p_2\left(\frac{1+q_1}{1-q_1q_2}\right) \Rightarrow \frac{p_1p_2(1+q_1)}{1-q_1q_2}$$

(iv) Since probability of scoring is $\frac{1}{3}$, expected number of kicks = $\frac{1}{\frac{1}{3}} = 3$.

Because one more kick is allowed after scoring, expected number = 3 + 1 = 4.

Example 2: S2007 P2 Q8

On average 1 in 8 people in a particular community is left-handed and the rest are right-handed. A sample of people is chosen at random, one by one, until a left-handed person is obtained. Find the probability that the number of people in the sample is less than 12. [3]

Another sample is chosen at random, one by one, until at least one left-handed and at least one right-handed person have been obtained. The number of people in this sample is denoted by N, so that the sample consists of

either : $(N-1)$ left-handed people followed by one right-handed person

or : $(N-1)$ right-handed people followed by one left-handed person.

Obtain $P(N = r)$, for $r \geqslant 2$, and show that $E(N) = 8\frac{1}{7}$. [5]

Let L be number of people chosen to obtain a left-handed person, L ~ Geo (0.125)

$$P(L<12) = 1-(0.875)^{11} = 0.770$$

Let R be number of people chosen to obtain a right-handed person, R ~ Geo (0.875)

$$P(r-1 \text{ left-handed and a right-handed}) = P(R = r-1) = \left(\frac{7}{8}\right)^{r-1}\left(\frac{1}{8}\right)$$

$$P(r-1 \text{ right-handed and a left-handed}) = P(L = N-1) = \left(\frac{7}{8}\right)\left(\frac{1}{8}\right)^{r-1}$$

Thus $P(N = r) = P(R = r-1) + P(L = r-1) = \left(\frac{7}{8}\right)\left(\frac{1}{8}\right)^{r-1} + \left(\frac{7}{8}\right)^{r-1}\left(\frac{1}{8}\right)$

Since $P(N = r) = P(R = r-1) + P(L = r-1)$

$$E(N) = E(L) + E(R) - P(L=1) - P(R=1)$$

$$= \frac{1}{\frac{1}{8}} + \frac{1}{\frac{7}{8}} - 1 = 8 + \frac{8}{7} - 1 = 8\frac{1}{7}$$

*Note: L = R = 1 is removed because $r = 1$ is not possible for N.

Example 3: W2002 P2 Q10 OR

An author sends his first manuscript to a large number of publishers, C, D, E, …, in turn, only approaching each one, after the first, if the one before has refused it. There is a constant probability $\frac{1}{4}$ that his manuscript will be accepted by each publisher approached. The random variable M is the number of publishers approached, up to and including the one who accepts the manuscript. Write down the values of $E(M)$ and $Var(M)$, and calculate

$$P\big(|M - E(M)| < \sqrt{\{Var(M)\}}\big).$$ [5]

For his second manuscript the author decides to approach two other publishers, A then B, for each of whom the probability of acceptance is $\frac{1}{2}$, before he approaches C, D, E, …. The probability of acceptance by each of these publishers remains $\frac{1}{4}$. The random variable N is the number of A, B, C, D, E, … approached, up to and including the one who accepts this manuscript. Write down the first few terms of the series for $E(M)$ and for $E(N)$. By comparing corresponding terms, after the second, and using your value for $E(M)$, show that $E(N) = \frac{5}{2}$. [6]

Use a similar method to show that $E(N^2) = \frac{27}{2}$, and hence find $Var(N)$. [3]

$M \sim Geo(0.25)$

$$E(M) = \frac{1}{0.25} = 4$$

$$Var(M) = \frac{0.75}{0.25^2} = 12$$

$$|M - E(M)| < \sqrt{Var(M)}$$

$$(M-4)^2 < 12$$

$$M^2 - 8M + 4 < 0$$

Consider $M^2 - 8M + 4 = 0$,

$$M = \frac{8 \pm \sqrt{64-16}}{2} = 7.464, -0.536$$

Thus $-0.536 < M < 7.464$

$\rightarrow M \leq 7$, since $M \in N$.

$P(M \leq 7)$

$= 1 - 0.75^7$

$= 0.867$

$E(M) = 1(0.25) + 2(0.75)(0.25) + 3(0.75^2)(0.25) + 4(0.75^3)(0.25) + ...$

$E(N) = 1(0.5) + 2(0.5^2) + 3(0.5^2)(0.25) + 4(0.5^2)(0.75)(0.25) + ...$

Comparing values after the second term,

$E(M) = 1(0.25) + 2(0.75)(0.25) + 3(0.75^2)(0.25) + 4(0.75^3)(0.25) + ...$

$E(M) - 0.625 = (0.75^2)[3(0.25) + 4(0.75)(0.25) + ...]$

$\frac{16}{9}E(M) - \frac{10}{9} = 3(0.25) + 4(0.75)(0.25) + ...$

$E(N) - 1 = (0.5^2)[3(0.25) + 4(0.75)(0.25) + ...]$

$4E(N) - 4 = 3(0.25) + 4(0.75)(0.25) + ...$

$4E(N) - 4 = \frac{16}{9}(4) - \frac{10}{9}$

$36E(N) - 36 = 64 - 10$

$36E(N) = 90 \rightarrow E(N) = \frac{90}{36} = \frac{5}{2}$

$E(M^2) = Var(M) + [E(M)]^2 = 12 + 16 = 28$

Doing the same to the M^2 and N^2 terms,

$E(M^2) = 1(0.25) + 4(0.75)(0.25) + 9(0.75^2)(0.25) + 16(0.75^3)(0.25) + ...$

$\frac{16}{9}[E(M^2) - 1] = 9(0.25) + 16(0.25) + ...$

$E(N^2) = 1(0.5) + 4(0.5^2) + 9(0.5^2)(0.25) + 16(0.5^2)(0.75)(0.25) + ...$

$4[E(N^2) - 1.5] = 9(0.25) + 16(0.25) + ...$

$= \frac{16}{9}[E(M^2) - 1]$

$36E(N^2) - 54 = 16(28 - 1)$

$36E(N^2) = 486 \rightarrow E(N^2) = \frac{27}{2}$

$Var(N) = E(N^2) - [E(N)]^2 = 13.5 - 6.25 = 7.25$

15.4 Negative Exponential Distribution

- LO: understand conditions under which a negative exponential distribution may be a suitable probability model;
- LO: recall and use the formula for the calculation of negative exponential probabilities;
- LO: recall and use the means and variances of a geometric distribution and negative exponential distribution.

❖ A continuous random variable, probability density function is $f(t) = \lambda e^{-\lambda t}, t \geq 0$

❖ Cumulative distribution function is $F(T) = \int_0^t \lambda e^{-\lambda t} dt = 1 - e^{-\lambda t}$

❖ $P(T > t) = 1 - F(T) = e^{-\lambda t}$

❖ $E(T) = \int_0^\infty t \cdot \lambda e^{-\lambda t} dt = \frac{1}{\lambda}$

- $Var(T) = \int_0^\infty t^2 \cdot \lambda e^{-\lambda t} dt - [E(T)]^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$
- All properties (area is 1, quartiles) also apply as it is a continuous random variable
- Conditions to use it follows a Poisson distribution (events must occur singly, randomly, at a constant mean rate and are independent)
- Some special properties of the exponential distribution:

No.	Property	Explanation
1	Memory less	$\Rightarrow P[(X > a+b) \mid (X > a)] = P(X > b)$ $\Rightarrow$ Can be proven using conditional probability (see exp. 3)
2	Is the continuous form of the Poisson Distribution	$\Rightarrow$ If X is the number of cars arriving in one minute, and if T is the time, in minutes, between successive cars, $X \sim Po(\lambda) \rightarrow f(t) = \lambda e^{-\lambda t}$

Example 1: S2002 P2 Q11 OR

A country bus driver picks up passengers randomly and independently at a mean rate of 12 per hour.

(i) Find, correct to 3 decimal places, the probability that he picks up at least 3 passengers in a period of 15 minutes. [2]

(ii) Show that the probability that he picks up at least one passenger in a period of t hours is

$$1 - e^{-12t}.$$ [2]

The time at which he picks up his first passenger is T hours. Explain why

$$P(T < t) = 1 - e^{-12t} \text{ for } t > 0.$$ [3]

Hence find the probability density function of T, and find the values of $E(T)$ and $Var(T)$. [4]

Find also the median time at which the first passenger is picked up. [3]

(i) Let X_t be the no. of passengers picked up in t hours $\rightarrow X_t \sim Po(12t)$. Thus $X_{0.25} \sim Po(3)$

$$P(\text{At least 3 passengers in 15 mins}) = 1 - P(X_{0.25} < 3)$$
$$= 1 - e^{-3}\left(\frac{3^0}{0!} + \frac{3^1}{1!} + \frac{3^2}{2!}\right)$$
$$= 1 - 8.5e^{-3} = 0.577$$

(ii) $P(\text{At least 1 passengers in t hours}) = 1 - P(X_t = 0)$

$$= 1 - e^{-12t}\left(\frac{12^0}{0!}\right) = 1 - e^{-12t}$$

$$P(\text{First passenger in t hours}) = P(T < t)$$
$$= P(\text{at least one passenger in t hours})$$
$$= 1 - e^{-12t}, t \geq 0$$

$$f(t) = \frac{d}{dt}[P(T < t)] = 12e^{-12t}, t \geq 0$$

$$E(T) = \frac{1}{12}, Var(T) = \frac{1}{144}$$

Let the median be m. Thus $F(m) = 0.5$;

$$1 - e^{-12m} = 0.5$$
$$-12m = \ln 0.5$$
$$m = \frac{1}{12}\ln 2 = 0.0578$$

Example 2: W2007 P2 Q9

Log-ons to a certain computer website occur randomly at a uniform average rate of 2.4 per minute. State the distribution of the number N of log-ons that occur during a period of t minutes. [1]

Obtain the probability that at least one log-on occurs during a period of t minutes. [2]

Hence obtain the probability density function of T, where T minutes is the interval between successive log-ons. [3]

Identify the distribution of T and state its mean and variance. [3]

$N \sim Po(2.4t)$ $\quad P(\text{at least one log-on in t minutes}) = 1 - P(N = 0)$

$$= 1 - e^{-2.4t}\left(\frac{[2.4t]^0}{0!}\right) = 1 - e^{-2.4t}$$

$$P(T < t) = F(T)$$
$$= P(\text{at least one log-on in t minutes})$$
$$= 1 - e^{-2.4t}$$

$$f(t) = \frac{d}{dt}[F(T)]$$
$$= 2.4e^{-2.4t}$$

Distribution function of T: negative exponential distribution

$$E(T) = \frac{1}{2.4} = \frac{5}{12} \qquad Var(T) = \frac{1}{2.4^2} = \frac{25}{144}$$

Example 3: S2005 P2 Q8

The continuous random variable X has probability density function given by

$$f(x) = \begin{cases} \frac{1}{\theta}e^{-\frac{x}{\theta}} & x \geqslant 0, \\ 0 & \text{otherwise,} \end{cases}$$

where θ is a positive constant. Show that the mean of X is θ, and determine which is the greater, $P(X < \theta)$ or $P(X > \theta)$. [4]

Given that a and b are positive, show that $P(X > a + b \mid X > a) = P(X > b)$. [3]

Explain the meaning of this equation in the case when X is the number of days after treatment for a particular disease when a patient needs further treatment. [2]

$$E(X) = \int_0^\infty \frac{x}{\theta} e^{-\frac{x}{\theta}} dx$$
$$= \left[-xe^{-\frac{x}{\theta}}\right]_0^\infty + \int_0^\infty e^{-\frac{x}{\theta}} dx$$
$$= \left[-\theta e^{-\frac{x}{\theta}}\right]_0^\infty$$
$$= \theta$$

$$P(X < \theta) = \int_0^\theta \frac{1}{\theta} e^{-\frac{x}{\theta}} dx$$
$$= \left[-e^{-\frac{x}{\theta}}\right]_0^\theta$$
$$= 1 - e^{-1}$$
$$= 0.632 > 0.5$$

Thus $P(X < \theta) > P(X > \theta)$.

$$P(X > a + b \mid X > a)$$
$$= \frac{P(X > a + b \cap X > a)}{P(X > a)}$$
$$= \frac{P(X > a + b)}{P(X > a)}$$
$$= \frac{e^{-\frac{a+b}{\theta}}}{e^{-\frac{a}{\theta}}}$$
$$= e^{\frac{-a-b+a}{\theta}}$$
$$= e^{-\frac{b}{\theta}} = P(X > b)$$

Meaning: The probability that a patient needs *b* further days to treat his disease given that he already received *a* days of previous treatment is equal to the probability that the patient requiring b days to treat the same illness at the beginning of the treatment.

CHAPTER 16: INFERENCE USING NORMAL AND t-DISTRIBUTIONS

16.1 Pre-Requisite Knowledge:

A) Distributions (Normal, Poisson, Binomial in P6 and P7)

B) Sampling and Estimation (P7)

1. Terms: sample versus population (sample is part of a population

2. Distribution of the sample mean

$$\text{If } X \sim N(\mu,\sigma^2), \text{ then } \overline{X} \sim N\left(\mu,\frac{\sigma^2}{n}\right) \text{ for all n.}$$

If X is not normally distributed, then $\overline{X} \sim N\left(\mu,\frac{\sigma^2}{n}\right)$ if $n \geq 30$. (Central Limit Theorem)

3. Unbiased Estimates

$$\text{Mean}: \hat{\mu} = \bar{x} = \frac{\sum x}{n}.$$

$$\text{Variance: } \hat{\sigma}^2 = \frac{n}{n-1}s^2 = \frac{1}{n-1}\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)$$

4. Confidence Intervals

$\left(\hat{\mu} \pm x\frac{\hat{\sigma}}{\sqrt{n}}\right)$, x is value obtained from t-table / z-table etc. (usually found to 2 d.p.)

Half-width of the interval: $x\frac{\hat{\sigma}}{\sqrt{n}}$ Interval width: $2x\frac{\hat{\sigma}}{\sqrt{n}}$

Meaning of a 98% confidence interval of paint volume as (700, 750): The probability that the population mean of the volume of paint used is between 700 *ml* and 750 *ml* is 98%.

5. Confidence interval for population proportion: n > 30,

$$X \sim B(n, p_s) \rightarrow \overline{X} \sim N \sim (np_s, p_sq_s)$$

Since $p_s = \frac{\overline{X}}{n}$, $p_s \sim N\left(p_s, \frac{p_sq_s}{n}\right)$

Thus a CI is $\left(p_s \pm x\sqrt{\frac{p_sq_s}{n}}\right)$. Same rules apply as from (4).

Example 1: Sampling (S2008 Mathematics P7 Q2)

The lengths of time people take to complete a certain type of puzzle are normally distributed with mean 48.8 minutes and standard deviation 15.6 minutes. The random variable X represents the time taken in minutes by a randomly chosen person to solve this type of puzzle. The times taken by random samples of 5 people are noted. The mean time $\bar{X}$ is calculated for each sample.

(i) State the distribution of $\bar{X}$, giving the values of any parameters. [2]

(ii) Find $P(\bar{X} < 50)$. [3]

(i) Since X is normally distributed, so is $\bar{X}$.

$$\bar{X} \sim N\left(48.8, \frac{15.6^2}{5}\right) = (48.8, 48.672)$$

(ii)

$$P(\bar{X} < 50) = P\left(Z < \frac{50 - 48.8}{\sqrt{48.672}}\right)$$
$$= P(Z < 0.172) \approx 0.568$$

Example 2: Sampling and Estimation (S2007 Mathematics P7 Q6)

The daily takings, $\$x$, for a shop were noted on 30 randomly chosen days. The takings are summarised by $\Sigma x = 31\,500$, $\Sigma x^2 = 33\,141\,816$.

(i) Calculate unbiased estimates of the population mean and variance of the shop's daily takings. [3]

(ii) Calculate a 98% confidence interval for the mean daily takings. [3]

The mean daily takings for a random sample of n days is found.

(iii) Estimate the value of n for which it is approximately 95% certain that the sample mean does not differ from the population mean by more than \$6. [3]

(i)

$$\hat{\mu} = \frac{31500}{30} = 1050$$

$$\widehat{\sigma^2} = \frac{1}{29}\left(33141816 - \frac{31500^2}{30}\right)$$
$$= 2304$$

(ii) For a 98% CI, use $z(2.326) = 0.99$.

Half width = $2.326\sqrt{\frac{2304}{30}} = 20.38$

A 98% CI is $(1050 \pm 20.38) = (1030, 1070)$.

(iii) Half width of interval is 6. For a 95% confidence interval, use $z(1.960) = 0.975$.

$$1.96\sqrt{\frac{2304}{n}} = 6 \Rightarrow n = \left(\frac{1.96(48)}{6}\right)^2 \approx 246$$

Example 3: Confidence Intervals (W2008 Mathematics P7 Q4)

Diameters of golf balls are known to be normally distributed with mean μ cm and standard deviation σ cm. A random sample of 130 golf balls was taken and the diameters, x cm, were measured. The results are summarised by $\Sigma x = 555.1$ and $\Sigma x^2 = 2371.30$.

(i) Calculate unbiased estimates of μ and σ^2. [3]

(ii) Calculate a 97% confidence interval for μ. [3]

(iii) 300 random samples of 130 balls are taken and a 97% confidence interval is calculated for each sample. How many of these intervals would you expect **not** to contain μ? [1]

(i)

$$\hat{\mu} = \frac{555.1}{130} = 4.27$$

$$\widehat{\sigma^2} = \frac{1}{129}\left(2371.30 - \frac{555.1^2}{130}\right)$$
$$= 0.00793$$

(ii) For a 97% CI, use $z(2.17) = 0.985$.

Half width = $2.17\sqrt{\frac{0.00793}{130}} = 0.0169$

A 97% CI is $(4.27 \pm 0.0169) = (4.25, 4.29)$.

(iii) 97% of the time μ is in the interval, thus

No. of intervals without $\mu = 300 \times 3\% = 9$.

C) Hypothesis Testing (P7)

The Language of Hypothesis Testing:

1. Null Hypothesis / H_0: Hypothesis which the researcher tries to disprove, reject or nullify.

2. Alternative Hypothesis / H_1: Hypothesis that is contrary to H_0. It is usually taken to be that the observations are the result of a real effect (with some amount of chance variation superposed).

3. Significance Level: Probability that the event observed is unlikely to happen.

4. Rejection Region / critical region: cut-off point where observations are unlikely to occur / criteria to reject H_0 – usually obtained from z or t tables

5. Test statistic: value used by calculation to make a conclusion for the test

6. Conclusion: reject H_0 if test statistic lies in critical region, accept if it does not.

7. one tailed upper-tail test: used to test for an **increase** in population mean / proportion

8. one tailed lower-tail test: used to test for a **decrease** in population mean / proportion

9. two-tailed test: used to test for a **change** in population mean / proportion

10. Type I error: Error made when H_0 is rejected when it is true

11. Type II error: Error made when H_0 is accepted when it is false

Typical Procedures for a Hypothesis Test

Step 1: Define variable, state assumptions if needed, calculate unknown sample mean / variance

Step 2: State hypothesis (Usually, for H_0: $\mu = x$, for H_1: $\mu > x$ / $\mu < x$ / $\mu \neq x$)

State 3: State criteria for rejecting H_0 (calculate degrees of freedom if t-distribution is used)

Step 4: Calculate test statistic

Step 5: State conclusion

Example 4: Hypothesis Testing: Binomial + Errors (W2003 Mathematics P7 Q6)

(i) Explain what is meant by

(a) a Type I error, [1]

(b) a Type II error. [1]

(ii) Roger thinks that a box contains 6 screws and 94 nails. Felix thinks that the box contains 30 screws and 70 nails. In order to test these assumptions they decide to take 5 items at random from the box and inspect them, replacing each item after it has been inspected, and accept Roger's hypothesis (the null hypothesis) if all 5 items are nails.

(a) Calculate the probability of a Type I error. [4]

(b) If Felix's hypothesis (the alternative hypothesis) is true, calculate the probability of a Type II error. [3]

(i) refer to above

(ii)(a) $P(\text{Type I error}) = P(< 5 \text{ nails}, p = 0.94) = 1 - 0.94^5 = 0.266$

(ii)(b) $P(\text{Type II error}) = P(5 \text{ nails}, p = 0.7) = 0.7^5 = 0.168$

Example 5: Hypothesis Testing: one sample z-test (W2008 Mathematics P7 Q2)

The times taken for the pupils in Ming's year group to do their English homework have a normal distribution with standard deviation 15.7 minutes. A teacher estimates that the mean time is 42 minutes. The times taken by a random sample of 3 students from the year group were 27, 35 and 43 minutes. Carry out a hypothesis test at the 10% significance level to determine whether the teacher's estimate for the mean should be accepted, stating the null and alternative hypotheses. [5]

Let X be the time taken for pupils to do their English homework, $X \sim N\left(\mu, 15.7^2\right)$.

Step 1: Since mean is unknown, $\hat{\mu} = \frac{27+35+43}{3} = 35$.

Step 2: State hypothesis: H_0: $\mu = 42$ (mean is 42 minutes), H_1: $\mu \neq 42$ (mean is not 42 minutes)

If H_0 is true, then $\overline{X} \sim N\left(42, 15.7^2\right)$.

Step 3: Critical Region

Conducting a two-tailed test, at 10% significance level, reject H_0 if $|z| > 1.645$. [$z_{(0.95)}$]

Step 4: Test statistic: $Z = \frac{35-42}{15.7} = -0.446$.

Step 5: Conclusion

Since $|Z| < 1.645$, do not reject H_0. There is not enough evidence, at 10% significance level that the mean of Z is not 42. The teacher's estimate can be accepted.

16.2 One Sample t-Test

- LO: formulate hypotheses and apply a hypothesis test concerning the population mean using a small sample drawn from a normal population of unknown variance, using a *t*-test;
- LO: determine a confidence interval for a population mean, based on a small sample from a normal population with unknown variance, using a *t*-distribution;

  - ❖ This is similar to P7 one sample z-test, but used when the sample size is normally distributed, small (≤ 30) and when population variance is unknown.
  - ❖ In a t-test, degrees of freedom, ν is used to obtain values from the t-table.
  - ❖ For a one-sample t-test, degrees of freedom, $\nu = \text{no.of samples} - 1$
  - ❖ Required assumption for this test: Sample given follows a normal distribution / Random samples are taken.
  - ❖ Test statistic, $T = \dfrac{X - \mu}{\dfrac{\sigma}{\sqrt{n}}}$

Example 1: W2003 P2 Q6

A random sample of 10 observations of a normal variable X has mean $\bar{x}$, where

$$\bar{x} = 4.344, \quad \sum (x - \bar{x})^2 = 0.8022.$$

Carry out a 2-tail test, at the 5% level of significance, to test whether the mean of X is 4.58. State your null and alternative hypotheses clearly. [7]

Step 1: Since variance is unknown, $\hat{\sigma}^2 = \dfrac{10}{9}\left(\dfrac{0.8022}{10}\right) = \dfrac{0.8022}{9}$.

Step 2: State hypothesis: $H_0: \mu = 4.58$, $H_1: \mu \neq 4.58$

If H_0 is true, then $\overline{X} \sim N\left(4.58, \dfrac{\hat{\sigma}^2}{10}\right)$.

Step 3: Critical Region

$\nu = 10 - 1 = 9$. Conducting a two-tailed test, Using t(9) at 5% significance level, reject H_0 if $|t| > 2.262$. [$t_{(9,\ 0.975)}$]

Step 4: Test statistic: $T = \dfrac{4.344 - 4.58}{\sqrt{\dfrac{0.8022}{90}}} = -2.50$.

Step 5: Conclusion

Since $|T| > 2.262$, reject H_0 in favour on H_1. There is enough evidence, at 5% significance level that the mean of X is not 4.58.

Example 2: S2011 P21 Q7

A greengrocer claims that his cabbages have a mean mass of more than 1.2 kg. In order to check his claim, he weighs 10 cabbages, chosen at random from his stock. The masses, in kg, are as follows.

1.26 1.24 1.17 1.23 1.18 1.25 1.19 1.20 1.21 1.18

Stating any assumption that you make, test at the 10% significance level whether the greengrocer's claim is supported by this evidence. [8]

Assumptions: Mass of the greengrocer's cabbages follows a normal distribution.

Let X be the mass of the greengrocer's cabbages in kg, $X \sim N(\mu, \sigma^2)$.

From a calculator, $\Sigma x = 12.11, \Sigma x^2 = 14.6745$

$$\hat{\mu} = \frac{\Sigma x}{n} = \frac{12.1}{10} = 1.211,\ \hat{\sigma}^2 = \frac{1}{n-1}\left(\Sigma x^2 - \frac{(\Sigma x)^2}{n}\right) = \frac{1}{9}\left(14.6745 - \frac{12.11^2}{10}\right) = \frac{929}{900000}$$

$H_0: \mu = 1.2, \qquad H_1: \mu > 1.2$

If H_0 is true, then $\overline{X} \sim N\left(1.2, \frac{929}{9000000}\right)$.

$\nu = 10 - 1 = 9$. Conducting a one-tail (upper-tail) test, using t(9) at 10% significance level, reject H_0 if $t > 1.383$. $[t_{(9,\ 0.9)}]$

Test statistic, $T = \frac{1.211 - 1.2}{\sqrt{\frac{929}{9000000}}} = 1.08$

Since $T < 1.083$, do not reject H_0. There is not enough evidence, at 10% significance level, that the mean mass of the cabbages is not 1.2kg. The greengrocer's claim is not justified.

Example 3: W2008 P2 Q7

The pulse rate of each member of a random sample of 25 adult UK males who exercise for a given period each week is measured in beats per minute. A 98% confidence interval for the mean pulse rate, μ beats per minute, for all such UK males was calculated as $61.21 < \mu < 64.39$, based on a *t*-distribution.

(i) Calculate the sample mean pulse rate and the standard deviation used in the calculation. [5]

(ii) State an assumption necessary for the validity of the confidence interval. [1]

(iii) The mean pulse rate for all UK males is 72 beats per minute. State, giving a reason, if it can be concluded that, on average, UK males who exercise have a reduced pulse rate. [2]

(i) $\hat{\mu} = \text{Midpoint of confidence interval}$

$$= \frac{61.21 + 64.39}{2} = 62.8$$

$\nu = 25 - 1 = 24$

t-value used = $t_{(24,\ 0.99)} = 2.492$

Half-width of interval = $\frac{64.39 - 61.21}{2} = 1.59$

$$2.492\left(\frac{\sigma}{\sqrt{25}}\right) = 1.59$$

$$\sigma = \frac{1.59(5)}{2.492} = 3.19\ (3\text{ s.f.})$$

(ii) The pulse rate of the members follows a normal distribution.

(iii) Yes, because 72 exceeds the upper limit of the confidence interval.

Example 4: W2009 P2 Q6

A machine produces metal discs whose diameters have a normal distribution. The mean of this distribution is intended to be 10 cm. Accuracy is checked by measuring the diameters of a random sample of six discs. The diameters, in cm, are as follows.

10.03 10.02 9.98 10.06 10.08 10.01

Calculate a 99% confidence interval for the mean diameter of all discs produced by the machine. [5]

Deduce a 99% confidence interval for the mean circumference of all discs produced by the machine. [1]

$$\hat{\mu} = \frac{\Sigma x}{n} = \frac{60.18}{6} = 10.03$$

$\nu = 6 - 1 = 5$

t-value used = $t_{(5,\ 0.995)} = 4.032$

$\hat{\sigma}^2 = \frac{6}{5}\left(\frac{603.6118}{6} - 10.03^2\right) = 0.00128$ Half-width = $4.032\left(\sqrt{\frac{0.00128}{6}}\right) = 0.0589$

Thus a 99% CI for the mean diameter:

$(10.03 \pm 0.0589) = (9.97, 10.09)$

A 99% CI for mean circumference = $\pi(9.97, 10.09) = (31.32, 31.70) \Rightarrow (31.3, 31.7)$

Note: The sample mean is used instead of the population mean (which is only involved if a hypothesis test is required).

16.3 Paired Sample t-Test

- LO: formulate hypotheses concerning the difference of population means, and apply, as appropriate, a paired sample *t*-test.
- LO: determine a confidence interval for a difference of population means, using a *t*-distribution as appropriate.

- ❖ Used to compare difference between two samples of identical sizes
- ❖ Usually involves a small sample size
- ❖ Involves:

(1) Two measures of the same subjects ('before' and 'after')
(2) Two separate samples but subjects are individually matched

- ❖ Conducted like a one-sample t-test, but the difference is the two samples are used to make a 'difference' sample.
- ❖ Test statistic is thus calculated using the new random variable D, where $D = \frac{d - \mu}{\frac{\sigma}{\sqrt{n}}}$.
- ❖ Since it is similar to a one-sample t-test, it is required to assume that the sample follows a normal distribution / sample chosen is random.

Example 1: S2009 P2 Q8

Part of a research study of identical twins who had been separated at birth involved a random sample of 9 pairs, in which one twin had been raised by the natural parents and the other by adoptive parents. The IQ scores of these twins were measured, with the following results.

Twin pair	1	2	3	4	5	6	7	8	9
IQ of twin raised by natural parents	82	92	115	132	88	95	112	83	123
IQ of twin raised by adoptive parents	92	88	115	134	97	94	107	88	130

It may be assumed that the difference in IQ scores has a normal distribution. The mean IQ scores of separated twins raised by natural parents and by adoptive parents are denoted by μ_N and μ_A respectively. Obtain a 90% confidence interval for $\mu_N - \mu_A$. [6]

One of the researchers claimed that there was no evidence of a difference between the two population means. State, giving a reason, whether the confidence interval supports this claim. [2]

We start by constructing a table for differences:

Twin pair	1	2	3	4	5	6	7	8	9
IQ of twin raised by natural parents	82	92	115	132	88	95	112	83	123
IQ of twin raised by adoptive parents	92	88	115	134	97	94	107	88	130
Difference (Natural – Adoptive)	-10	4	0	-2	-9	1	5	-5	-7

From a calculator, $\Sigma d = -23, \Sigma d^2 = 301$

$\hat{\mu} = \frac{\Sigma d}{n} = \frac{-23}{9}$

$\hat{\sigma}^2 = \frac{9}{8}\left[\frac{301}{9} - \left(\frac{23}{9}\right)^2\right] = \frac{545}{18}$

$\nu = 9 - 1 = 8$.

Thus use $t_{(8,\ 0.95)} = 1.860$.

Half-width of interval: $1.860\left(\sqrt{\frac{545}{162}}\right) = 3.412$

Thus a 95% CI is $\left(-\frac{23}{9} \pm 3.412\right) = (-5.97, 0.86)$

Since 0 lies in the confidence interval, the professor's claim is supported.

Example 2: S2011 P23 Q8

A company decides that its employees should follow an exercise programme for 30 minutes each day, with the aim that they lose weight and increase productivity. The weights, in kg, of a random sample of 8 employees at the start of the programme and after following the programme for 6 weeks are shown in the table.

Employee	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*
Weight before (kg)	98.6	87.3	90.4	85.2	100.5	92.4	89.9	91.3
Weight after (kg)	93.5	85.2	88.2	84.6	95.4	89.3	86.0	87.6

Assuming that loss in weight is normally distributed, find a 95% confidence interval for the mean loss in weight of the company's employees. [6]

Test at the 5% significance level whether, after the exercise programme, there is a reduction of more than 2.5 kg in the population mean weight. [5]

Employee	A	B	C	D	E	F	G	H
Weight before (kg)	98.6	87.3	90.4	85.2	100.5	92.4	89.9	91.3
Weight after (kg)	93.5	85.2	88.2	84.6	95.4	89.3	86.0	87.6
Difference (B – A)	5.1	2.1	2.2	0.6	5.1	3.1	3.9	3.7

Let *D* be the difference in weight before and after the programme in kg, $D \sim N(\mu, \sigma^2)$.

From a calculator, $\Sigma d = 25.8, \Sigma d^2 = 100.14$

$\hat{\mu} = \frac{\Sigma d}{n} = \frac{25.8}{8} = 3.225$

$\hat{\sigma}^2 = \frac{1}{7}\left(100.14 - \frac{25.8^2}{8}\right) = \frac{3387}{1400}$

$\nu = 8 - 1 = 7$.

Thus use $t_{(7,\ 0.975)} = 2.365$.

Half-width of interval: $2.365\left(\sqrt{\frac{3387}{11200}}\right) = 1.301$

Thus a 95% CI is $(3.225 \pm 1.301) = (1.92, 4.53)$

H_0: $\mu_D = 2.5$, H_1: $\mu_D > 2.5$

If H_0 is true, then $\overline{D} \sim N\left(2.5, \frac{3387}{11200}\right)$.

$\nu = 8 - 1 = 7$. Conducting a one-tail (upper-tail) test, using t(7) at 5% significance level, reject H_0 if t > 1.895. $[t_{(7,\ 0.95)}]$

Test statistic, $D = \frac{3.225 - 2.5}{\sqrt{\frac{3387}{11200}}} = 1.32$

Since D < 1.832, do not reject H_0. There is not enough evidence, at 5% significance level, that there is a reduction of more than 2.5 kg in mean weight.

Example 3: S2010 P23 Q10

Carpal Tunnel syndrome is a condition which affects a person's ability to grip with their hands. Researchers tested a treatment for this syndrome which was applied to 8 randomly chosen patients. A pre-treatment and a post-treatment test of grip was given to each patient, with the following results, measured in kg.

Patient	1	2	3	4	5	6	7	8
Pre-treatment grip	24.3	29.5	28.0	28.5	21.5	28.7	25.1	26.3
Post-treatment grip	28.3	34.6	30.3	31.6	21.5	29.8	26.0	27.5

Stating any required assumption, test, at the 1% significance level, whether the mean grip of people with the syndrome increases after undergoing the treatment. [9]

It is given that there is evidence at the 10% significance level that the mean grip increases by more than w kg. Find an inequality for w. [3]

Patient	1	2	3	4	5	6	7	8
Pre-treatment grip	24.3	29.5	28.0	28.5	21.5	28.7	25.1	26.3
Post-treatment grip	28.3	34.6	30.3	31.6	21.5	29.8	26.0	27.5
Difference (Post – Pre)	4.0	5.1	2.3	3.1	0	1.1	0.9	1.2

Assumption: Population of the difference of the grip of the patients follow a normal distribution.

Let P be the difference in weight before and after the programme in kg, $P \sim N\left(\mu, \sigma^2\right)$.

From a calculator, $\Sigma p = 17.7, \Sigma p^2 = 60.37$

$$\Rightarrow \hat{\mu} = \frac{\Sigma p}{n} = \frac{17.7}{8} = 2.2125, \quad \hat{\sigma}^2 = \frac{1}{7}\left(60.37 - \frac{17.7^2}{8}\right) = \frac{16967}{5600}$$

H_0: $\mu_P = 0$, H_1: $\mu_P > 0$

If H_0 is true, then $\bar{P} \sim 0\left(0, \frac{16967}{44800}\right)$.

$\nu = 8 - 1 = 7$. Conducting a one-tail (upper-tail) test, using t(7) at 1% significance level, reject H_0 if t > 2.998. [$t_{(7,\ 0.99)}$]

Test statistic, $D = \frac{2.2125}{\sqrt{\frac{16967}{44800}}} = 3.59$

Since D > 2.998, reject H_0 in favour of H_1. There is enough evidence, at 1% significance level, that mean grip of the patients increased after the treatment.

t-value used: $t_{(7,\ 0.9)}$ = 1.415 Because results provide evidence,

$$\frac{2.2125 - w}{\sqrt{\frac{16967}{44800}}} > 1.415 \rightarrow 2.2125 - w > 0.8708 \Rightarrow w < 1.34$$

16.4 Two Sample Tests

- LO: calculate a pooled estimate of a population variance from two samples (calculations based on either raw or summarised data may be required);
- LO: formulate hypotheses concerning the difference of population means, and apply, as appropriate:
  - a 2-sample *t*-test,
  - a test using a normal distribution,

  (the ability to select the test appropriate to the circumstances of a problem is expected);
- LO: determine a confidence interval for a difference of population means, using a *t*-distribution, or a normal distribution, as appropriate.
  - ❖ Usually used to test a difference in mean between random samples
  - ❖ Test used depends on sample size (z if big, t if small)

Z-Tests:

- ❖ Test statistic follows linear combinations of random variables. Since sample size > 30, $\overline{Z_1} \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1}\right)$ and $\overline{Z_2} \sim N\left(\mu_2, \frac{\sigma_2^2}{n_2}\right)$, $\overline{Z_1} - \overline{Z_2} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$
- ❖ Thus test statistic $Z = \frac{x_1 - x_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$.
- ❖ General procedures follows a one-sample z-test

Example 1: W2009 P2 Q8

150 sheep, chosen from a large flock of sheep, were divided into two groups of 75. Over a fixed period, one group had their grazing controlled and the other group grazed freely. The gains in weight, in kg, were recorded for each animal and the table below shows the sample means and the unbiased estimates of the population variances for the two samples.

	Sample mean	Unbiased estimate of population variance
Controlled grazing	19.14	20.54
Free grazing	15.36	9.84

It is required to test whether the population mean for sheep having their grazing controlled exceeds the population mean for sheep grazing freely by less than 5 kg. State, giving a reason, if it is necessary for the validity of the test to assume that the two population variances are equal. [1]

Stating any other assumption, carry out the test at the 5% significance level. [8]

No. since sample size given is large (≥ 30), Central Limit Theorem can be used.

Assumptions: Sample chosen must be random.

Step 1: Define random variable.

Let C be mass of sheep in controlled grazing group, $\overline{C} \sim N\left(19.14, \frac{20.54}{75}\right)$

Let F be mass of sheep in controlled grazing group, $\overline{F} \sim N\left(15.36, \frac{9.84}{75}\right)$

Step 2: State hypothesis H_0: $\mu_c - \mu_f = 5$, H_1: $\mu_c - \mu_f < 5$

If H_0 is true, then $\overline{C} - \overline{F} \sim N\left(5, \frac{20.54}{75} + \frac{9.84}{75}\right) \rightarrow \overline{C} - \overline{F} \sim N(5, 0.40507)$

Step 3: Critical region

Using a one-tail (lower-tail) test at the 5% significance level, reject H_0 if $z < -1.645$.

Step 4: Test statistic, $Z = \frac{19.14 - 15.36 - 5}{\sqrt{0.40507}} = -1.92$

Step 5: Conclusion: Since $Z < -1.645$, reject H_0 in favour of H_1. There is enough evidence, at the 5% significance level that the population mass of the sheep having their grazing controlled exceeds the population of the sheep grazing freely by less than 5 kg.

Example 2: W2008 P2 Q11 OR

A perfume manufacturer had one bottle-filling machine, but because of increased sales a second machine was obtained. In order to compare the performance of the two machines, a random sample of 50 bottles filled by the first machine and a random sample of 60 bottles filled by the second machine were checked. The volumes of the contents from the first machine, x_1 ml, and from the second machine, x_2 ml, are summarised by

$$\Sigma x_1 = 1492.0, \quad \Sigma x_1^2 = 44\,529.52, \quad \Sigma x_2 = 1803.6, \quad \Sigma x_2^2 = 54\,220.84.$$

The volumes have distributions with means μ_1 ml and μ_2 ml for the first and second machines respectively. Test, at the 2% significance level, whether μ_2 is greater than μ_1. [10]

Find the set of values of α for which there would be evidence at the α% significance level that $\mu_2 - \mu_1 > 0.1$. [4]

Let X be the difference in mean volume of the two perfume samples.

$$\hat{\mu}_1 = \frac{1492.0}{50} = 29.84, \quad \hat{\mu}_2 = \frac{1803.6}{60} = 30.06$$

$$\hat{\sigma}_1^2 = \frac{1}{49}\left(44529.52 - \frac{1492^2}{50}\right) = \frac{206}{1225}, \quad \hat{\sigma}_2^2 = \frac{1}{59}\left(54220.84 - \frac{1803.6^2}{60}\right) = \frac{578}{7375}$$

H_0: $\mu_2 - \mu_1 = 0$, H_1: $\mu_2 - \mu_1 > 0$

If H_0 is true, then $\overline{X} \sim N\left(0, \frac{103}{30625} + \frac{289}{221250}\right) \rightarrow \overline{X} \sim N(0, 0.00467)$

Using a one-tail (upper tail) test, at the 2% significance level, reject H_0 if $z > 2.0545$.

Test statistic, $Z = \frac{30.06 - 29.84}{\sqrt{0.00467}} = 3.22$

Since $Z > 2.0545$, reject H_0 in favour of H_1. There is enough evidence, at the 2% significance level that μ_2 is greater than μ_1.

Note: If $\mu_1 - \mu_2$ is used then a lower-tail test would have to be used.

New test statistic: $Z = \frac{0.22 - 0.1}{\sqrt{0.00467}} = 1.756$

For evidence to be present, tabulated value < 1.756. Tabulated value obtained is 0.9608,

Thus $\alpha_{\min} = 100(1 - 0.9608) = 3.92, \alpha \geq 3.9/4.0$.

Example 3: W2006 P2 Q8

Two types of cloth, Type A and Type B, were tested for flame-resistance. 45 pieces of Type A and 48 pieces of Type B, of equal size, were selected at random. Each was set alight in the same way and the times to extinction of the flames were measured. The results, x_A minutes and x_B minutes, are summarised by

$$\Sigma x_A = 276.5, \quad \Sigma x_A^2 = 1823.0, \quad \Sigma x_B = 267.8, \quad \Sigma x_B^2 = 1649.2.$$

Test, at the 2% significance level, whether there is a difference in the mean times to extinction for the two types of cloth. [8]

Explain, in the context of the question, the meaning of 'at the 2% significance level'. [1]

$$\hat{\mu}_A = \frac{276.5}{45} = 6.1444, \quad \hat{\mu}_B = \frac{267.8}{48} = 5.5792$$

$$\hat{\sigma}_A^2 = \frac{1}{44}\left(1823 - \frac{276.5^2}{45}\right) = 2.81957, \quad \hat{\sigma}_B^2 = \frac{1}{47}\left(1649.2 - \frac{267.8^2}{48}\right) = 3.3000$$

Let A be time used to extinct flames by type A cloth, $\overline{A} \sim N\left(6.1444, \frac{2.81957}{45}\right)$

Let B be time used to extinct flames by type B cloth, $\overline{B} \sim N\left(5.5792, \frac{3.3000}{48}\right)$

H_0: $\mu_A - \mu_B = 0$, H_1: $\mu_A - \mu_B \neq 0$

If H_0 is true, then $\overline{X} \sim N\left(0, \frac{2.81957}{45} + \frac{3.3000}{48}\right) \rightarrow \overline{X} \sim N(0, 0.13141)$

Using a two-tailed test, at the 2% significance level, reject H_0 if $|z| > 2.326$.

Test statistic, $Z = \frac{6.14444 - 5.5792}{\sqrt{0.13141}} = 1.56$

Since $|Z| < 2.0545$, do not reject H_0. There is not enough evidence, at the 2% significance level that there is a difference in mean times to the extinction of the cloth.

Meaning of the 2% significance level: Probability of concluding that the probability of the sample means to not be equal when they are in fact equal is 2%. (Type I error)

T-Tests:

- Used when total sample size is less than 30.
- Since sample size is small and population variance is unknown, new assumptions are needed: Populations follow a normal distribution and random samples are taken, and both populations have a common variance.
- Since variance is assumed to be same, a new pooled estimate of the variance is needed, and is $\hat{\sigma}^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2} = \frac{(n_1 - 1)\hat{\sigma}_1^2 + (n_2 - 1)\hat{\sigma}_2^2}{n_1 + n_2 - 2}$.
- Degrees of freedom, $\nu = n_1 + n_2 - 2$
- Test statistic, $T = \frac{x_1 - x_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\hat{\sigma}^2}{n_1} + \frac{\hat{\sigma}^2}{n_2}}}$
- General procedures are similar to one-sample t-test in Section 16.2.

Example 1: S2004 P2 Q12 OR

The numbers of road accidents per week in a certain town form an approximate normal distribution. A random sample of size 7 is

$$9,\quad 5,\quad 8,\quad 6,\quad 11,\quad 13,\quad 8.$$

Find a 95% confidence interval for the mean number of accidents per week. [5]

In a country area, a random sample of size 4 gives the numbers of road accidents per week as

$$7,\quad 3,\quad 10,\quad 2.$$

Assuming that both distributions are normal with a common variance, obtain a pooled estimate of that variance. [4]

Test, at the 5% level of significance, whether, on average, accidents occur more frequently in the town than in the country. [5]

$\mu_1 = \frac{60}{7}, \quad \hat{\sigma}_1^2 = \frac{1}{6}\left(560 - \frac{3600}{7}\right) = \frac{160}{21}$

$\nu = 7-1 = 6$, thus use $t_{(6,\,0.975)} = 2.447$

Half-width = $2.447\sqrt{\frac{160}{147}} = 2.553$

Thus a 95% CI for is $\left(\frac{60}{7} \pm 2.553 = 6.02, 11.12\right)$

$\mu_2 = \frac{22}{4} = 5.5, \quad \hat{\sigma}_2^2 = \frac{1}{3}\left(162 - \frac{22^2}{4}\right) = \frac{41}{3}$

$\hat{\sigma}^2 = \frac{6\left(\frac{160}{21}\right) + 3\left(\frac{41}{3}\right)}{7+4-2} = \frac{607}{63} \approx 9.635$

$H_0: \mu_1 - \mu_2 = 0, \quad H_1: \mu_1 - \mu_2 > 0$

If H_0 is true, then $\overline{X} \sim N\left(0, \frac{9.635}{7} + \frac{9.635}{4}\right) \rightarrow \overline{X} \sim N(0,\ 3.785)$.

$\nu = 7+4-2 = 9$. Conducting a one-tail (upper-tail) test, using t(9) at 5% significance level, reject H_0 if t > 1.833. $[t_{(9,\,0.95)}]$

Test statistic, $T = \frac{\frac{60}{7} - 5.5}{\sqrt{3.785}} = 1.58$

Since T > 1.833, do not reject H_0. There is not enough evidence, at 5% significance level that accidents occur more frequently in the town than in the country.

Example 2: S2010 P21 Q8

An examination involved writing an essay. In order to compare the time taken to write the essay by students in two large colleges, a sample of 12 students from college A and a sample of 8 students from college B were randomly selected. The times, t_A and t_B, taken for these students to write the essay were measured, correct to the nearest minute, and are summarised by

$$n_A = 12, \quad \Sigma t_A = 257, \quad \Sigma t_A^2 = 5629, \quad n_B = 8, \quad \Sigma t_B = 206, \quad \Sigma t_B^2 = 5359.$$

Stating any required assumptions, calculate a 95% confidence interval for the difference in the population means. [8]

State, giving a reason, whether your confidence interval supports the statement that the population means, for the two colleges, are equal. [1]

Assumption 1: Time taken for students from both colleges is normally distributed.

Assumption 2: Population variance of the time taken to write the essay for both schools i equal.

$\hat{\mu}_A = \frac{257}{12} = 21.4167, \quad \hat{\mu}_B = \frac{206}{8} = 25.75 \; s_A^2 = \frac{5629}{12} - \left(\frac{257}{12}\right)^2 = \frac{1499}{144}, \quad s_B^2 = \frac{5359}{8} - 25.75^2 = \frac{109}{16}$

Thus a pooled estimate is $\hat{\sigma}^2 = \dfrac{12(\frac{1499}{144}) + 8(\frac{109}{16})}{12+8-2} = \dfrac{2513}{216}$

$\nu = 8+12-2 = 18$. Thus use $t_{(18,\ 0.975)} = 2.101$

Half width of interval: $2.101\sqrt{\dfrac{2513}{2592} + \dfrac{2513}{1728}} = 3.271$

Thus a 95% CI is $(25.75 - 21.4167 \pm 3.271) = (1.06, 7.60)$

The confidence interval does not contain 0, so the relationship is not supported.

Note: $(21.4167 - 25.75 \pm 3.271) = (-7.60, -1.06)$ is also an equally valid answer.

Example 3: S2002 P2 Q10

A racing driver claims that he can bring his car to rest from a speed of $150\,\text{km h}^{-1}$ in a shorter distance with radial-ply tyres than with other types of tyre. As evidence for this claim, he lists the stopping distances on a race track in twelve random trials, six with radial-ply tyres, and six with other types of tyre.

	Stopping distances (m)					
Radial-ply tyres	123	164	139	161	118	135
Other types of tyre	156	137	134	173	155	151

Test at the 5% significance level whether the racing driver's claim is justified, stating your null hypothesis and your alternative hypothesis clearly. Assume that the populations sampled are normal and have equal variances. [10]

Find a 95% confidence interval for the difference between the population means for radial-ply tyres and other types of tyre. [3]

$\hat{\mu}_R = \dfrac{840}{6} = 140,\ \ \hat{\mu}_O = \dfrac{906}{6} = 151\ \ s_R^2 = \dfrac{119416}{6} - 140^2 = 302.67,\ \ s_O^2 = \dfrac{137816}{6} - 151^2 = 168.33$

Thus a pooled estimate is $\hat{\sigma}^2 = \dfrac{6(302.67) + 6(168.33)}{6+6-2} = 282.6$

Let R be stopping distances of radial-ply tyres, $\overline{R} \sim N(140, 47.1)$

Let O be stopping distances of other tyres, $\overline{O} \sim N(151, 47.1)$

H_0: $\mu_R - \mu_O = 0$ (No difference between radial-ply and other tyres)

H_1: $\mu_R - \mu_O < 0$ (Radial-ply requires a less distance to stop)

If H_0 is true, then $\overline{R} - \overline{O} \sim N(0,\ 94.2)$.

$\nu = 6+6-2 = 10$. Conducting a one-tail (lower-tail) test, using t(10) at 5% significance level, reject H_0 if $t < -1.812$. $[t_{(10,\ 0.95)}]$

Test statistic, $T = \dfrac{140-151}{\sqrt{94.2}} = -1.13$

Since $T > -1.812$, do not reject H_0. There is not enough evidence, at 5% significance level that the radial-ply tyres take a less distance to stop. The racing driver's claim is not justified.

Tabulated t-value is $t_{(10,\ 0.975)} = 2.228$

Half-width is $2.228\sqrt{94.2} = 21.62$, Thus a 95% CI is $(140 - 151 \pm 21.62) = (-32.6, 10.6)$

Note: A paired sampled test is not used here because different samples are involved.

CHAPTER 17: χ^2 TESTS

- Pre-Requisite Knowledge: All previous Distributions and Hypothesis Testing
- The χ^2 distribution is the distribution of the sum of squared Z values
- The χ^2 significance test has one parameter, ν.
- The shape is different for different values of ν, just like the t-distribution
- The test is conducted as a one-tailed (upper tail) test
- The critical value/region can be found from the χ^2 table
- Test statistic: $\chi^2 = \sum \frac{(O-E)^2}{E}$
- O is observed data: it is given in the question
- E is expected data: it is obtained by calculation

17.1 Test for Goodness of Fit

- LO: fit a theoretical distribution, as prescribed by a given hypothesis, to given data (questions will not involve lengthy calculations);
- LO: use a χ^2 test, with the appropriate number of degrees of freedom, to carry out the corresponding goodness of fit analysis (classes should be combined so that each expected frequency is at least 5);

- To test how well a distribution fits a particular data
- Exam questions may involve all previous distributions learnt (given ratio, binomial, normal, poisson, continuous random variables, geometric and exponential distributions)
- Steps in a χ^2 goodness of fit test:

1. Define variable, find mean / variance if unknown

2. State hypothesis

3. Calculate expected frequency using $nP(X = x)$. Combine consecutive rows (until all expected frequencies are at least 5)

4. Calculate Degrees of freedom, $\nu = \text{No. of classes} - \text{No. of restrictions}$

No. of classes (rows/ columns after combining) No. of restrictions (Usually 1 for total frequency known, 1 for estimating mean, 1 for estimating variance)

5. State condition for rejecting H_0 (using table of χ^2 values)

6. Calculate test statistic

7. State conclusion

Example 1: Binomial – S2006 P2 Q10

Torch batteries of a particular brand are sold in packs of 3. A store receives a large delivery of packs. The manager selects a random sample of 60 packs and tests all of the batteries. The numbers of faulty batteries in the packs are summarised in the following table.

Number faulty	0	1	2	3
Number of packs	28	28	4	0

(i) State one assumption about the batteries in a pack which you would need to make for a binomial distribution to be appropriate. [1]

(ii) Fit a binomial distribution to the data and carry out a goodness of fit test at the 10% significance level. [8]

(iii) 10 packs are selected at random from the delivery. Estimate the probability that at least two of these packs contain 3 faulty batteries. [3]

(i) The probability of picking a faulty battery is the same in all packs. / The probability of picking a faulty battery is independent of the pack taken or other batteries.

(ii) Step 1 : Define the Variable,

Let X be the number of faulty batteries in packs of 3.

$$np = \frac{28 + 2(4)}{60} = 0.6 \Rightarrow 3p = 0.6,\ p = 0.2$$

Step 2: Stating Hypothesis

H_0: X~B(3, 0.2) [Assume distribution is true], H_1: It is not true that X~ B(3, 0.2).

Step 3: Calculate Expected Frequency

x	Observed Data	P ($X = x$)	Expected Data [60 P($X=x$)]
0	28	${}^3C_0(0.8)^3 = 0.512$	60 (0.512) = 30.72
1	28	${}^3C_1(0.2)(0.8)^2 = 0.384$	60 (0.384) = 23.04
2	4	${}^3C_2(0.2)^2(0.8) = 0.096$	60 (0.096) = 5.76
3	0	${}^3C_3(0.2)^3 = 0.008$	60 (0.008) = 0.48

Combining the last two rows,

x	Observed Data	Expected Data
0	28	30.72
1	28	23.04
2 or more	4	6.24

Step 4: Degrees of Freedom

$$\nu = \text{No. of classes} - \text{No. of restrictions} = 3 - 2 = 1$$

-2: total frequency is known, mean is estimated

Step 5: State Criteria for Rejecting H_0

Using $x^2(1)$ at 10% significance level, reject H_0 if $x^2 > 2.706 \rightarrow [1, 0.9]$ from table.

Step 6: Calculate Test Statistic

Test statistic, $\chi^2 = \sum \frac{(O-E)^2}{E}$

$$= \frac{2.72^2}{30.72} + \frac{5.96^2}{23.04} + \frac{2.24^2}{6.24}$$

$$= 2.587 \approx 2.59 \text{ (3s.f.)}$$

Step 7: Conclusion

Since $\chi^2 = 2.59 < 2.706$, Do not reject H_0. We conclude that X ~ B (3, 0.2) is and adequate model for the data.

(iii) $\text{P(3 faulty batteries)} = (0.2)^3 = 0.008$

Let Y be no. of packs with 3 faulty batteries out of 10 packs, Y ~ B (10, 0.008)

$$P(Y \geq 2)$$
$$= 1 - P(Y - 1) - P(Y = 0)$$
$$= 1 - 10(0.008)(0.992)^9 - (0.992)^{10}$$
$$= 0.00276 (3\text{s.f.})$$

Example 2: Poisson – S2003 P2 Q7

An official in a cycle race notes the number of cyclists passing her in each of 60 random one-minute intervals. Her results are given in the table.

Number of cyclists	0	1	2	3	4	5 or more
Frequency	10	12	14	16	8	0

Fit a Poisson distribution to the data, and test its goodness of fit at the 5% level of significance. [8]

$$\lambda = \frac{12 + 28 + 48 + 32}{60} = 2$$

Let X be the number of cyclist passing in one minute.

H_0: X~ Po (2) H_1: It is not true that X ~ Po (2).

x	Observed Data	P ($X = x$)	Expected Data [60P($X=x$)]
0	10	$e^{-2} = 0.13534$	60 (0.13534) = 8.120
1	12	$2e^{-2} = 0.27067$	60 (0.27067) = 16.24
2	14	$2e^{-2} = 0.27067$	60 (0.27067) = 16.24
3	16	$\frac{8}{6}e^{-2} = 0.18045$	60(0.18045) = 10.83
4	8	$\frac{16}{24}e^{-2} = 0.09022$	60(0.09022) = 5.413
≥ 5	0	$1 - P(X \leq 4) = 0.05265$	60(0.05265) = 3.159

Combining the last two rows,

x	0	1	2	3	4 or more
Observed Data	10	12	14	16	8
Expected Data	8.120	16.24	16.24	10.83	8.572

$\nu = \text{No. of classes} - \text{No. of restrictions} = 5 - 2 = 3$

Using $x^2(3)$ at 5% significance level, reject H_0 if $x^2 > 7.815 \rightarrow [3, 0.95]$ from table.

$$\text{Test statistic}, \chi^2 = \sum \frac{(O-E)^2}{E}$$
$$= \frac{1.88^2}{8.12} + \frac{4.24^2}{16.24} + \frac{2.24^2}{16.24} + \frac{5.17^2}{10.83} + \frac{0.572^2}{8.572}$$
$$= 4.357 \approx 4.36 \text{ (3s.f.)}$$

Since $\chi^2 = 4.36 < 7.815$, Do not reject H_0. We conclude that X ~ Po (2) is an adequate model for the data.

Example 3: S2007 P2 Q11 OR (Given Ratio)

A large population consists of four species, A, B, C and D, which are claimed to be in the proportions p_1, p_2, p_3 and p_4 respectively, where $\sum_{i=1}^{4} p_i = 1$. In order to test this claim a random sample of size n is selected and the numbers of each species are found to be a, b, c and d respectively. Show that the χ^2 statistic for a goodness of fit test simplifies to

$$\chi^2 = \frac{a^2}{np_1} + \frac{b^2}{np_2} + \frac{c^2}{np_3} + \frac{d^2}{np_4} - n. \qquad [4]$$

It is given that $p_1 = p_2$, $p_3 = p_4$ and the values of a, b, c and d are 15, 25, 39 and 21 respectively. Denoting the common value of p_1 and p_2 by p, show that

$$\chi^2 = \frac{8.5}{p} + \frac{19.62}{0.5 - p} - 100. \qquad [2]$$

Determine the value p_0 of p for which this value of χ^2 is a minimum. [3]

Carry out the goodness of fit test at the 5% significance level, with $p = p_0$. [4]

State, giving your reason, the conclusion of the test for values of p other than p_0. [1]

	A	B	C	D	Total
Expected	np_1	np_2	np_3	np_4	n
Observed	a	b	c	d	n

$$\text{Test statistic, } \chi^2 = \sum \frac{(O-E)^2}{E}$$
$$= \frac{(a-np_1)^2}{np_1} + \frac{(b-np_2)^2}{np_2} + \frac{(c-np_3)^2}{np_3} + \frac{(d-np_4)^2}{np_4}$$
$$= \frac{a^2}{np_1} - 2a + np_1 + \frac{b^2}{np_2} - 2b + np_2 + \frac{c^2}{np_3} - 2c + np_3 + \frac{d^2}{np_4} - 2d + np_4$$
$$= \frac{a^2}{np_1} + \frac{b^2}{np_2} + \frac{c^2}{np_3} + \frac{d^2}{np_4} - 2n + n$$
$$= \frac{a^2}{np_1} + \frac{b^2}{np_2} + \frac{c^2}{np_3} + \frac{d^2}{np_4} - n$$

$$\chi^2 = \frac{a^2}{np_1} + \frac{b^2}{np_2} + \frac{c^2}{np_3} + \frac{d^2}{np_4} - n$$
$$= \frac{225}{100p} + \frac{625}{100p} + \frac{1521}{100(0.5-p)} + \frac{441}{100(0.5-p)} - 100 \qquad a+b+c+d=100,$$
$$p_1 = p_2 = p, p_3 = p_4 = 0.5-p,$$
$$= \frac{8.5}{p} + \frac{19.62}{0.5-p} - 100 \qquad \text{since } p+p+0.5-p+0.5-p=1.$$

$$\frac{d\chi^2}{dp} = \frac{19.62}{(0.5-p)^2} - \frac{8.5}{p^2} \qquad \frac{d^2\chi^2}{dp^2} = \frac{17}{p^3} + \frac{39.24}{(0.5-p)^3}$$

$\frac{d\chi^2}{dp} = 0$ when $p = p_0$, When $p = p_0, \frac{d^2\chi^2}{dp^2} = 3605 > 0$

$$\frac{19.62}{(0.5-p_0)^2} = \frac{8.5}{p_0^2} \qquad \text{Thus } p_0 = 0.198 \text{ (3 s.f.)}$$

$$11.12p_0^2 + 8.5p_0 - 2.125 = 0$$

$$p_0 = 0.19847, -0.9639 (R)$$

H_0: A: B: C: D = 0.198: 0.198: 0.302: 0.302.

H_1: It is not true that A: B: C: D = 0.198: 0.198: 0.302: 0.302.

$\nu = \text{No. of classes} - \text{No. of restrictions} = 4 - 1 = 3$

Using $x^2(3)$ at 5% significance level, reject H_0 if $x^2 > 7.815 \rightarrow [3, 0.95]$ from table.

When $p_0 = 0.198$, $\chi^2 = \frac{8.5}{0.198} + \frac{19.62}{0.302} - 100 = 7.896 \approx 7.90$

Since $\chi^2 > 7.815$, Reject H_0 in favour of H_1. Thus, we always conclude that it is not true that A: B: C: D = 0.198: 0.198: 0.302: 0.302.

Since other values of p_0 always gives a greater test statistic value, H_0 is always rejected, and thus we always conclude that it is not true that A: B: C: D = 0.198: 0.198: 0.302: 0.302.

Summary (Degrees of freedom used)

distribution	condition	v
Uniform / Given Ratio		n – 1
Binomial	p known p unknown, $\bar{x}$ = np	n – 1 n – 2
Poisson	λ known λ unknown, $\bar{x}$ = λ	n – 1 n – 2
Normal	μ and σ^2 known μ and σ^2 unknown	n – 1 n – 3

17.2 Test for Independence

- LO: use a χ^2 test, with the appropriate number of degrees of freedom, for independence in a contingency table (Yates' correction is not required, but classes should be combined so that each expected frequency is at least 5).

❖ To test for dependence / relationship between variables
❖ Classes cannot be combined (expected frequency must be larger than 5)
❖ Steps in a χ^2 independence test:

1. State hypothesis

2. Calculate expected frequency using a contingency table.

$$E = \frac{\Sigma\, row \times \Sigma\, column}{\Sigma\, total}$$

Example: Given the table below:

Car colour \ Income	Low	Medium	High	Total
Red	A	B	C	80
Blue	D	E	F	120
Total	60	100	40	200

$A = 60 \times 80 \div 200 = 24$
$B = 100 \times 80 \div 200 = 40$
$C = 40 \times 80 \div 200 = 16$
$D = 60 \times 120 \div 200 = 36$
$E = 100 \times 120 \div 200 = 60$
$F = 40 \times 120 \div 200 = 24$

Notice that the total of each row and column must add up to the total of each row / column.

3. Calculate Degrees of freedom, $\nu = (\text{No. of rows} - 1)(\text{No. of columns} - 1)$

4. State condition for rejecting H_0 (using table of χ^2 values)

5. Calculate test statistic

6. State conclusion

Example 1: W2004 P2 Q8

Three suppliers, P, Q and R, are under examination for the quality of their items, which are rated poor, fair or good. A random sample of items is taken from each supplier, and the numbers in each category are given in the table.

Supplier	Poor	Fair	Good
P	17	40	123
Q	12	17	31
R	11	43	66

Test, at the 5% significance level, whether quality rating is independent of supplier. [8]

Step 1: Hypothesis

H_0: Quality rating is independent of supplier. (H_0 assumes independence)

H_1: Quality rating is dependent of supplier.

Step 2: Make a contingency table for expected values

Supplier	**Poor**	**Fair**	**Good**	**Total**
P	**20**	**50**	**110**	**180**
Q	**6.67**	**16.67**	**36.67**	**60**
R	**13.33**	**33.33**	**73.33**	**120**
Total	**40**	**100**	**220**	**360**

Step 3: Degrees of Freedom

$$\nu = (\text{No. of rows} - 1)(\text{No. of columns} - 1) = (3-1)(3-1) = 2 \times 2 = 4$$

Step 4: State Criteria for Rejecting H_0

Using $x^2(4)$ at 5% significance level, reject H_0 if $x^2 > 9.488 \rightarrow [4, 0.95]$ from table.

Step 5: Calculate Test Statistic

$$\text{Test statistic, } \chi^2 = \sum \frac{(O-E)^2}{E}$$
$$= \frac{3^2}{20} + \frac{5.33^2}{6.67} + \frac{2.33^2}{13.33} + \frac{10^2}{50} + \frac{0.33^2}{16.33} + \frac{9.67^2}{33.33} + \frac{13^2}{110} + \frac{5.67^2}{36.67} + \frac{7.33^2}{73.33}$$
$$= 13.07 \approx 13.1 \text{ (3s.f.)}$$

Step 6: Conclusion

Since $\chi^2 = 13.1 > 9.488$, H_0 is rejected in favour of H_1. We conclude that quality rating is not independent of supplier.

Example 2: S2002 P2 Q8

A telephone engineer tests a random sample of mobile phones for accuracy in connecting to the desired number, and for sound quality when the number has been connected. He measures each property as Low, Medium or High. The numbers of phones in each category are given in the table.

		Sound quality		
		Low	Medium	High
Accuracy	Low	12	23	76
	Medium	8	12	28
	High	13	30	112

Test, at the 5% level of significance, whether sound quality is independent of accuracy. [6]

If the number of phones with Low accuracy and High sound quality had been 0 instead of 76 (and all the other numbers in the table were unchanged) what problem would you have encountered in your solution? [2]

H_0: Sound quality is independent of accuracy.

H_1: Sound quality is dependent of accuracy.

	Low	**Medium**	**High**	**Total**
Low	**11.666**	**22.978**	**76.357**	**111**
Medium	**5.0446**	**9.9363**	**33.019**	**48**
High	**16.290**	**32.086**	**106.62**	**155**
Total	**33**	**65**	**216**	**314**

$\nu = (3-1)(3-1) = 4$.

Using $x^2(4)$ at 5% significance level, reject H_0 if $x^2 > 9.488 \rightarrow [4, 0.95]$ from table.

Test statistic, $\chi^2 = \sum \frac{(O-E)^2}{E}$

$$= \frac{0.334^2}{11.666} + \frac{2.9554^2}{5.0446} + \frac{3.290^2}{16.290} + \frac{0.022^2}{22.978} + \frac{2.0637^2}{9.9363} + \frac{2.086^2}{32.086} + \frac{0.357^2}{76.357} + \frac{5.019^2}{33.019} + \frac{5.38^2}{106.62}$$

$= 4.01$(3s.f.)

Since $\chi^2 < 9.488$, do not reject H_0. We conclude that quality rating is independent of supplier.

If that column is changed, the new contingency table becomes:

	Low	**Medium**	**High**	**Total**
Low	4.8529	9.9559	20.588	**35**
Medium	6.6555	13.109	28.235	**48**
High	21.492	42.332	91.176	**155**
Total	**33**	**65**	**140**	**238**

Since an expected frequency is less than 5 (the Low/low cell), a x^2 test cannot be performed.

Example 3: W2007 P2 Q11 OR

An experiment was performed to investigate a possible association between a person's hair-colour and eye-colour. The researcher selected a random sample of 60 people and classified their hair and their eyes as either light or dark. 36 people had light hair and 25 had light eyes. A χ^2 test of independence is to be carried out at the 5% significance level. Denoting by n the number of people with both light hair and light eyes, find the set of values of n for which it would be accepted that there is an association between hair-colour and eye-colour. [11]

Another researcher performed the experiment with a random sample of 40 people. In this sample, of the 24 people with light eyes 7 had light hair. Of the people with light hair 3 had dark eyes. Explain why the use of a χ^2 test on this data would be inappropriate. [3]

Contingency table: Expected data

Eye \ Hair	Light	Dark	Total
Light	**15**	**10**	25
Dark	**21**	**14**	35
Total	36	24	60

Table: Observed data

Eye \ Hair	Light	Dark	Total
Light	n	$25-n$	25
Dark	$36-n$	$n-1$	35
Total	36	24	60

H_0: There is no association between hair colour and eye colour.

H_1: There is no association between hair colour and eye colour.

$v = (2-1)(2-1) = 1$. Using $x^2(1, 0.95)$ at 5% significance level, reject H_0 if $x^2 > 3.841$.

$$\text{Test statistic, } \chi^2 = \sum \frac{(O-E)^2}{E}$$

$$3.841 < \frac{(n-15)^2}{15} + \frac{(36-n-21)^2}{21} + \frac{(25-n-10)^2}{10} + \frac{(n-1-14)^2}{14}$$

$$\left(\frac{1}{15}+\frac{1}{21}+\frac{1}{10}+\frac{1}{14}\right)(n-15)^2 > 3.841$$

$$(n-15)^2 > 3.841\left(\frac{7}{2}\right) = 13.44$$

$$n-15 > \sqrt{13.44} \text{ or } n-15 < -\sqrt{13.44}$$

$$n > 18.666 \text{ or } n < 11.334$$

Since $n \geq 1$ and $n \leq 25$, the set of values are $\{n:\ 1 < n < 11,\ 19 < n < 25,\ n \in N\}$.

New observed data:

Eye \ Hair	Light	Dark	Total
Light	7	17	24
Dark	3	13	16
Total	10	30	40

New expected data:

Eye \ Hair	Light	Dark	Total
Light	**6**	**18**	24
Dark	**4**	**12**	16
Total	10	30	40

- The expected amount of people with light hair and dark eyes is 4, which is less than 5.
- This requires the rows to be combined and the degrees of freedom will be reduced to 0.
- Hence, the use of a chi-squared test on this set of data is inappropriate.

CHAPTER 18: BIVARIATE DATA

18.1 Correlation

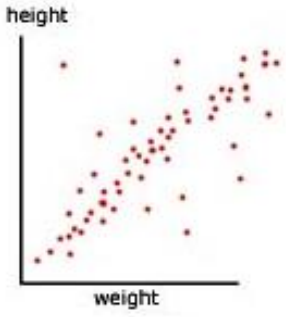
- LO: understand the concept of least squares, correlation and regression lines in the context of a scatter diagram;
- LO: calculate, both from simple raw data and from summarised data, the product moment correlation coefficient;
- LO: relate, in simple terms, the value of the product moment correlation coefficient to the appearance of the scatter diagram, with particular reference to the interpretation of cases where the value of the product moment correlation coefficient is close to +1, −1 or 0.
- LO: carry out a hypothesis test based on the product moment correlation coefficient.

A) Scatter Diagrams

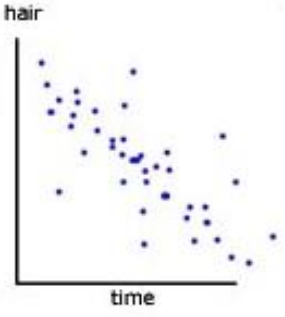
❖ Bivariate data: data connecting two variables, x and y. for example:

Independent variable (x)	**Dependent variable (y)**
diameter of the stem of a plant	average length of the leaf of the plant
a student's marks in a Math test	the same student's marks in a History test
the length of a spring	the extension of the spring when a 50g mass is placed

❖ Scatter Diagrams: graph produced when pairs of points in bivariate data are plotted together. This helps to deduce the relationship between the two variables.



General pattern: Height increases as weight increases (Positive linear correlation)



General pattern: Hair decreases as time increases (Negative linear correlation)

No general pattern between eye colour and age. (No correlation)

❖ Mathematically, a scatter diagram allows us to look for a relationship in the two variables in the form of $y = f(x)$. As we are looking for a linear correlation, the function $y = f(x)$ will be a straight line known as a regression line.

❖ We will use two different types of regression lines:

y on x: $y = a + bx$ (usually used, when x is the independent variable)
x on y: $x = c + dy$ (rarely used, when y is the independent variable)

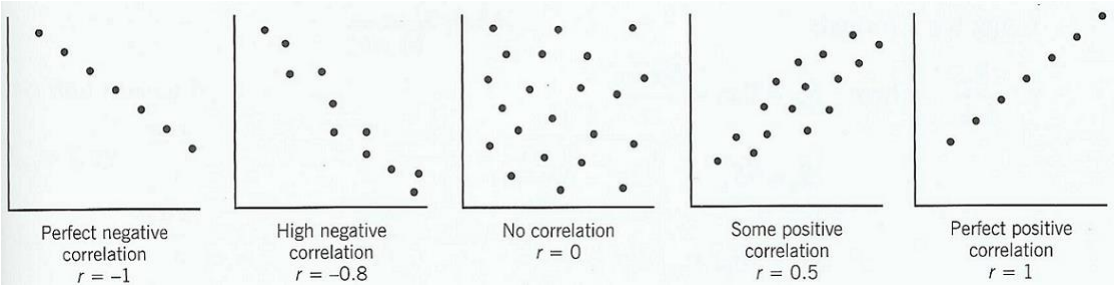
B) (Pearson's) Product Moment Correlation Coefficient (PMCC)

❖ The PMCC, r is a numerical value between 1 and –1 ($-1 \leq r \leq 1$).

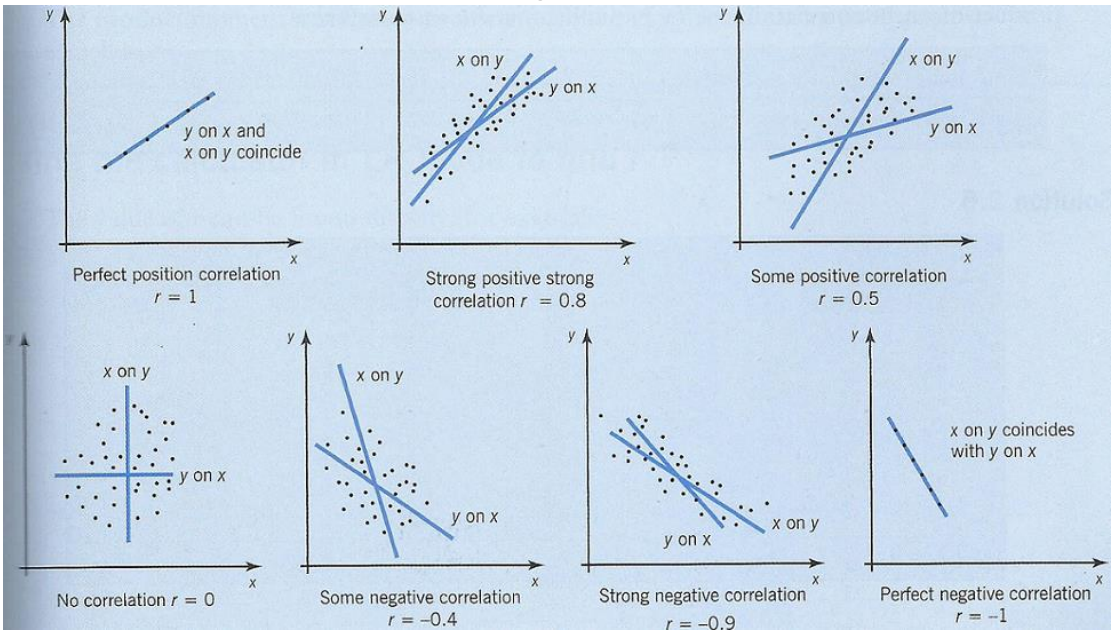
❖ It is an indicator of the linear degree of scatter of points in a scatter diagram. The nearer the value of r is to 1 or –1, the closer the points on the scatter diagram are on the regression line.

❖ When r approaches 1, the points will resemble a straight line with a positive gradient.

- ❖ When r approaches –1, the points will resemble a straight line with a positive gradient.
- ❖ When r approaches 0, there is no correlation or very weak correlation (the points will not behave like a straight line, but other relationships like $y = x^2$ or $xy = 0.01$ are possible)
- ❖ The PMCC is unaffected by coding e.g. changing x to $2x + 3$.



Plotting the two regression lines, y on x and x on y on a scatter diagram can also give a good idea in the value of r. The closer the two lines are together, the nearer r is to ± 1 as seen below:


y on x and x on y coincide
Perfect position correlation r = 1
x on y
y on x
Strong positive strong correlation r = 0.8
x on y
y on x
Some positive correlation r = 0.5
x on y
y on x
No correlation r = 0
x on y
y on x
Some negative correlation r = –0.4
y on x
x on y
Strong negative correlation r = –0.9
x on y coincides with y on x
Perfect negative correlation r = –1

Calculation of PMCC:

Step 1: Find $n, \bar{x}, \bar{y}, \sum x^2, \sum y^2$ and $\sum xy$ (omit this step if they are given / no raw data)

Step 2: Find the variance of x,

$$\boldsymbol{s_{xx} = \frac{\sum x^2}{n} - \bar{x}^2}$$

Step 3: Find the variance of y,

$$\boldsymbol{s_{yy} = \frac{\sum y^2}{n} - \bar{y}^2}$$

Step 4: Find the covariance of x and y,

$$\boldsymbol{s_{xy} = \frac{\sum xy}{n} - \bar{x}\bar{y}}$$

Step 5: Calculate r using

$$\boldsymbol{r = \frac{s_{xy}}{s_x s_y} = \frac{s_{xy}}{\sqrt{s_{xx} s_{yy}}}}$$

Note: Notations s_{xx}, s_{yy} and s_{xy} are used again to calculate regression lines in the next section.

P.S. Some books use the notation $S_{xy} = ns_{xy}$, $S_{xx} = ns_{xx}$ and $S_{yy} = ns_{yy}$.

Correlation versus causation: Two variables might correlate with one another but that does not mean that one event causes another to happen.

- For example, both the annual production of Coca-Cola and the amount of people owning computers both increased for the past few years but it would be foolish to conclude that events like drinking Cola-cola and owning a computer are related to one another.

C) Hypothesis Test for Population PMCC

- Exam questions will ask to test for behaviour of population PMCC, ρ , whether it is positive, negative or non-zero. The sample PMCC calculated, r, is used as test statistic.
- Typical Procedures for a Hypothesis Test:

Step 1: State hypothesis (H_0: $\rho = 0$, for H_1: $\rho > 0$ / $\rho < 0$ / $\rho \neq 0$)

State 2: State criteria for rejecting H_0 (use the table on MF 10 page 8 – there are different values for one or two-tailed test and is based on sample size – degrees of freedom are not required)

Step 3: State conclusion (Compare *r* with tabular value based on rejection criteria)

Example: S2010 P23 Q9

(a) The following are values of the product moment correlation coefficient between the x and y values of three different large samples of bivariate data. State what each indicates about the appearance of a scatter diagram illustrating the data.

(i) -1,

(ii) 0.02,

(iii) 0.92.

[3]

(b) In 1852 Dr William Farr published data on deaths due to cholera during an outbreak of the disease in London. The table shows the altitude (in feet, above the level of the river Thames) at which people lived and the corresponding number of deaths from cholera per 10 000 people.

Altitude, x	10	30	50	70	90	100	350
Number of deaths, y	102	65	34	27	22	17	8

[$\Sigma x = 700$, $\Sigma x^2 = 149\,000$, $\Sigma y = 275$, $\Sigma y^2 = 17\,351$, $\Sigma xy = 13\,040$.]

(i) Calculate the product moment correlation coefficient. [3]

(ii) Test, at the 5% significance level, whether there is evidence of negative correlation. [4]

(a)(i) The points when joined together form a straight line with negative gradient.

(ii) The points have almost no linear relation when joined together.

(iii) The points when joined together almost form a straight line with positive gradient.

(b)(i)

$$s_{xy} = \frac{13040}{7} - \frac{700(275)}{49} = -2066$$

$$s_{xx} = \frac{149000}{7} - 100^2 = 11286$$

$$s_{yy} = \frac{17351}{7} - \left(\frac{275}{7}\right)^2 = 935.35$$

$$r = \frac{s_{xy}}{s_x s_y} = \frac{-2066}{\sqrt{11286(935.35)}}$$

$$= -0.636 \ (3\ d.p.)$$

(ii)

H_0: $\rho = 0$

H_1: $\rho < 0$

Using a one-tailed (lower-tail) test at $n = 7$, at 5% significance level, reject H_0 if $\rho < -0.669$.

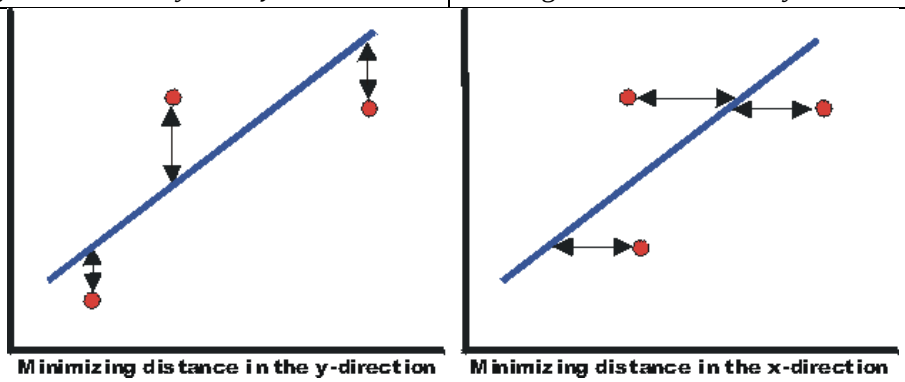
Since $r = -0.636 > -0.669$, do not reject H_0. At 5% significance level, there is not enough evidence for negative correlation.

18.2 Regression

- LO: calculate, both from simple raw data and from summarised data, the equations of regression lines, and appreciate the distinction between the regression line of y on x and that of x on y ;
- LO: select and use, in the context of a problem, the appropriate regression line to estimate a value, and understand the uncertainties associated with such estimations;
- LO: recall and use the facts that both regression lines pass through the mean centre $(\bar{x}, \bar{y})$ and the product moment correlation coefficient r and the regression coefficients b_1, b_2 are related by $r^2 = b_1 b_2$;

A) Regression Lines

- ❖ A regression line is a line that best describes the relationship between two sets of data.
- ❖ In AS Physics, the line is known as the line of best fit, and passes through as many points of possible.
- ❖ Now, we are able to calculate the equation of the lines we want to plot. The calculation techniques in the previous section allow us to calculate equations of two **least square regression lines**.

Regression line of y on x: $y = a + bx$	Regression line of x on y: $x = c + dy$
	
❖ As distance can be positive or negative in this context, a least squares regression line minimises the sum of squares of the distances instead. (similar to concept from χ^2 tests)	
❖ Minimises the sum of squares of vertical distance of points from the line ❖ Used when x is the independent variable and y is the dependent variable. ❖ y-intercept of the line is a ❖ $\boldsymbol{b = \text{gradient of line} = \frac{s_{xy}}{s_{xx}}}$	❖ Minimises the sum of squares of horizontal distance of points from the line ❖ Used when y is the independent variable and x is the dependent variable. ❖ y- intercept of the line is $-\frac{c}{d}$ ❖ $\boldsymbol{d = \frac{1}{\text{gradient of line}} = \frac{s_{xy}}{s_{yy}}}$

Properties of both regression lines, $y = a + bx$ and $x = c + dy$:

1. As $b = \frac{s_{xy}}{s_{xx}}$ and $d = \frac{s_{xy}}{s_{yy}}$,

$$bd = \frac{s_{xy}^2}{s_{xx} s_{yy}} = r^2 \rightarrow r = \pm\sqrt{bd}$$

If s_{xy} is positive, b, d and r are all positive, thus:

- ❖ If both $b > 0$ and $d > 0$, $r = \sqrt{bd}$;
- ❖ If both $b < 0$ and $d < 0$, $r = -\sqrt{bd}$.

2. They both pass through the point $(\bar{x},\ \bar{y})$
   - ❖ This fact is used to calculate the line equations, as $(\bar{x},\ \bar{y})$ is substituted to obtain *a* or *c*.

Steps on calculating the equation of a regression line:
1. Calculate $b = \frac{s_{xy}}{s_{xx}}$ (for *y* on *x* line) or $d = \frac{s_{xy}}{s_{yy}}$ (for *x* on *y* line)
2. Substitute $(\bar{x},\ \bar{y})$ to find *a* (for *y* on *x* line) or *c* (for *x* on *y* line)

B) Estimations using the Regression Line

- ❖ To make an estimation using a regression line, substitute *x* or *y* into the equation and obtain the corresponding value.
- ❖ An estimation using a regression line is only accurate if and only if:

1. No extrapolation of the line is required (the value used is in the range of values of data taken)
2. The value of PMCC is very high ($|r| \approx 1$).
   - ❖ The equation of *y* on *x*: $y = a + bx$ is used, when *x* is the independent variable
   - ❖ The equation of *x* on *y*: $x = c + dy$ is used when *y* is the independent variable

Note: If the value of PMCC is very high ($|r| \approx 1$), the choice of regression line does not matter.

Example 1: S2008 P2 Q10

The time in the air, *x* seconds, and the score, *y*, for 12 randomly chosen divers competing at a diving competition for a reverse half-somersault event are summarised as follows.

$$\Sigma x = 5.22 \quad \Sigma x^2 = 2.2774 \quad \Sigma y = 85.9 \quad \Sigma y^2 = 618.11 \quad \Sigma xy = 37.338$$

Obtain an estimate for the score of a diver who takes 0.40 seconds. [4]

Find the product moment correlation coefficient for the sample. [2]

Stating your hypotheses, test at the 5% significance level whether there is a non-zero correlation between the two variables. [4]

$$\bar{x} = \frac{5.22}{12} = 0.435,\ \ \bar{y} = \frac{85.9}{12} = 7.15833$$

$$s_{xx} = \frac{2.2774}{12} - 0.435^2 = 5.5833 \times 10^{-4}$$

$$s_{xy} = \frac{37.338}{12} - \bar{x}\bar{y} = -2.375 \times 10^{-3}$$

$$b = \frac{s_{xy}}{s_{xx}} = -4.254$$

At $(\bar{x},\ \bar{y})$:

$$a = 7.15833 + 4.254(0.435) = 9.01$$

The regression line is $y = 9.01 - 4.254x$.

$x = 0.4, y = 9.01 - 4.254(0.4) = 7.31\ s$

$$s_{yy} = \frac{618.11}{12} - 7.158333^2 = 0.2674$$

$$r = \frac{s_{xy}}{s_x s_y} = \frac{-2.375 \times 10^{-3}}{\sqrt{5.583 \times 10^{-4}(0.2674)}} = -0.194$$

H_0: $\rho = 0$

H_1: $\rho \neq 0$

Using a two-tailed test at $n = 12$, at 5% significance level, reject H_0 if $|\rho| > 0.576$.
Since $|r| = 0.194 < 0.576$, do not reject H_0.
At 5% significance level, there is not enough evidence for non-zero correlation.

Example 2: S2007 P2 Q7

Laboratory measurements were made of the speed of sound, v cm s^{-1}, in air at 9 different controlled temperatures, θ °C. The results are given in the following table.

θ	−10	−5	0	5	10	15	20	25	30
v	3.24	3.29	3.31	3.34	3.38	3.41	3.44	3.47	3.51

[$\Sigma\theta = 90$, $\Sigma\theta^2 = 2400$, $\Sigma v = 30.39$, $\Sigma v^2 = 102.6805$, $\Sigma\theta v = 313.65$.]

Determine whether a linear model is appropriate for the data. [3]

Obtain the equation of a suitable regression line. [3]

Estimate the temperature of the air when the speed of sound is 4 cm s^{-1} and comment on the reliability of this estimate. [2]

$$\bar{v} = \frac{30.39}{9} = 3.3767, \quad \bar{\theta} = \frac{90}{9} = 10$$

$$s_{v\theta} = \frac{313.65}{9} - (10)(3.3767) = 1.0833$$

$$s_{vv} = \frac{102.6805}{9} - 3.3767^2 = 7.064 \times 10^{-3}$$

$$s_{\theta\theta} = \frac{2400}{9} - 10^2 = \frac{500}{3}$$

$$r = \frac{1.0833}{\sqrt{\frac{500}{3}(7.064 \times 10^{-3})}} \approx 0.998 \ (3\ s.f.)$$

Regression line: $v = a + b\theta$

$$b = \frac{s_{v\theta}}{s_{\theta\theta}} = 6.46 \times 10^{-3}$$

At $(\bar{\theta}, \bar{v})$:

$$a = 3.3767 - 6.46 \times 10^{-2} = 3.31$$

$$v = 3.31 + 0.00646\,\theta$$

When $v = 4$,

$$\theta = \frac{4 - 3.31}{0.000646} \approx 106/107°C$$

Since 4 lies outside of the range of measurements taken for this laboratory experiment, the estimation is unreliable.

Problem Solving:

Example 1: W2004 P2 Q11 OR

A random sample of size n from a bivariate distribution is denoted by (x_r, y_r), $r = 1, 2, \ldots, n$. Show that, if the regression line of y on x passes through the origin of its scatter diagram, then

$$\bar{y}\sum_{r=1}^{n} x_r^2 = \bar{x}\sum_{r=1}^{n} x_r y_r, \qquad \text{(A)}$$

where $(\bar{x}, \bar{y})$ is the mean point of the sample. [4]

The table below gives a random sample of 5 pairs of values.

$x_1 = 0$	$x_2 = 4$	$x_3 = 2$	$x_4 = 1$	$x_5 = 3$
$y_1 = 3$	$y_2 = 12$	$y_3 = -1$	$y_4 = 2$	$y_5 = 14$

Verify that condition (A) is satisfied by this sample. [1]

Find the regression line, L, of y on x for the above sample. [3]

The equation $y = kx$ represents a line l through the origin of the scatter diagram of this sample, and $S = \sum_{r=1}^{5}(y_r - kx_r)^2$. Draw a diagram to show what S represents. [2]

By expressing S as a quadratic in k, show that, as k varies, S is a minimum when l is L. [4]

Regression line is $y = a + bx$

It passes through (0, 0) $\rightarrow a = 0$.

At $(\bar{x},\ \bar{y}), \bar{y} = b\bar{x}$

$$\bar{y}\, S_{xx} = \bar{x}\, S_{xy}$$

$$\bar{y}\left(\frac{\sum_{r=1}^{n} x_r^2}{n} - \bar{x}^2\right) = \bar{x}\left(\frac{\sum_{r=1}^{n} x_r y_r}{n} - \bar{x}\bar{y}\right)$$

$$\bar{y}\,\frac{\sum_{r=1}^{n} x_r^2}{n} - \bar{x}^2\bar{y} = \bar{x}\,\frac{\sum_{r=1}^{n} x_r y_r}{n} - \bar{x}^2\bar{y}$$

$$\bar{y}\,\frac{\sum_{r=1}^{n} x_r^2}{n} = \bar{x}\,\frac{\sum_{r=1}^{n} x_r y_r}{n}$$

$$\sum_{r=1}^{5} x_r^2 = 30, \sum_{r=1}^{5} x_r y_r = 90$$

$$\bar{x} = 2, \bar{y} = 6, n = 5$$

LHS: $\frac{6(30)}{5} = 36 = RHS$, condition is satisfied.

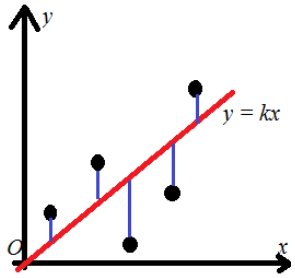
$$s_{xy} = \frac{90}{5} - 2(6) = 6, s_{xx} = \frac{30}{5} - 2^2 = 2$$

$$b = \frac{s_{xy}}{s_{xx}} = 3$$

At (2, 6), $a = 6 - 3(2) = 0$

The regression line is $y = 3x$.

S represents sum of squares of the vertical distance between the points and the line. (sum of length of blue lines squared)



$$S = 3^2 + (12 - 4k)^2 + (-1 - 2k)^2 + (2 - k)^2 + (14 - 3k)^2$$

$$= 30k^2 - 180k + 354$$

$$\frac{dS}{dk} = 60k - 180, \quad \frac{d^2S}{dk^2} = 60$$

When $k = 3, \frac{dS}{dk} = 0, \frac{d^2S}{dk^2} > 0$.

Thus S is minimum when $k = 3$,

S is minimum when l is L.

Example 2: W2005 P2 Q9

A teacher gives two tests, X and Y, to her class of n pupils, and labels their marks (x_1, y_1), (x_2, y_2), …, (x_n, y_n) respectively. An extra pupil is added to her class and the teacher gives her the provisional marks $(\bar{x}, \bar{y})$, the means of the marks already awarded. Show that the regression lines of y on x before and after the extra pupil's arrival are identical. [5]

x	24	38	41	57
y	34	21	63	22

Given that $n = 4$ and that the marks are as in the table, find the regression line of y on x. [2]

The extra pupil now sits test X and scores 46. Use the regression line to estimate her mark for test Y. Comment on the suitability of the method in this case, justifying your answer. [3]

Before arrival: $s_{xy} = \frac{\sum xy}{n} - \bar{x}\bar{y}$, $\quad s_{xx} = \frac{\sum x^2}{n} - \bar{x}^2$

After arrival: $s_{xy} = \frac{\sum xy + \bar{x}\bar{y}}{n+1} - \bar{x}\bar{y}$, $\quad s_{xx} = \frac{\sum x^2 + \bar{x}^2}{n+1} - \bar{x}^2$

$$Old\ b = \frac{s_{xy}}{s_{xx}}\ (old)$$

$$= \frac{\frac{\sum xy}{n} - \bar{x}\bar{y}}{\frac{\sum x^2}{n} - \bar{x}^2}$$

$$= \frac{\sum xy - n\bar{x}\bar{y}}{\sum x^2 - n\bar{x}^2}$$

$$New\ b = \frac{s_{xy}}{s_{xx}}\ (new)$$

$$= \frac{\frac{\sum xy + \bar{x}\bar{y}}{n+1} - \bar{x}\bar{y}}{\frac{\sum x^2 + \bar{x}^2}{n+1} - \bar{x}^2}$$

$$= \frac{\sum xy - (n+1)\bar{x}\bar{y} + \bar{x}\bar{y}}{\sum x^2 - (n+1)\bar{x}^2 + \bar{x}^2}$$

$$= \frac{\sum xy - n\bar{x}\bar{y}}{\sum x^2 - n\bar{x}^2} = \text{old b (SHOWN)}$$

Old line: $y = a_1 + bx$ At $(\bar{x}, \bar{y})$, $a_1 = \bar{y} - b\bar{x}$	New line: $y = a_2 + bx$ At $(\bar{x}, \bar{y})$, $a_2 = \bar{y} - b\bar{x}$, hence $a_1 = a_2$.

Since values of *a* and *b* for both equations are the same, the regression lines are identical.

From a calculator, $\bar{x} = 40$, $\bar{y} = 35$, $\sum x^2 = 6950$, $\sum y^2 = 6050$, $\sum xy = 5451$

Hence, $s_{xy} = -37.25$, $s_{xx} = 137.5$, $b = \frac{-37.25}{137.5} = -0.271$

At $(40, 35)$, $a = 35 + 0.271(40) = 45.8$. The equation is $y = 45.8 - 0.271x$.

When $x = 46$, $y = 45.8 - 0.271(46) = 33.8 \approx 34$ marks.

$S_{yy} = \frac{6050}{4} - 35^2 = 287.5$. Hence $r = \frac{S_{xy}}{S_x S_y} = \frac{-37.25}{\sqrt{287.5(137.5)}} = -0.187$

Since value of *r* is close to 0, the method is not suitable.

Example 3: W2011 P2 Q10 OR

The regression line of y on x obtained from a random sample of five pairs of values of x and y is

$$y = 2.5x - 1.5.$$

The data is given in the following table.

x	1	2	4	2	6
y	2	3	6	p	q

(i) Show that $p + q = 19$. [3]

(ii) Find the values of p and q. [6]

(iii) Determine the value of the product moment correlation coefficient for this sample. [3]

(iv) It is later discovered that the values of x given in the table have each been divided by 10 (that is, the actual values are 10, 20, 40, 20, 60). **Without any further calculation**, state

(a) the equation of the actual regression line of y on x,

(b) the value of the actual product moment correlation coefficient.

[2]

(i) The regression line must past through $(\bar{x}, \bar{y})$. $\bar{x} = 3$, $\bar{y} = \frac{11+p+q}{5}$

$$\frac{11 + p + q}{5} = 2.5(3) - 1.5$$
$$11 + p + q = 37.5 - 7.5 \rightarrow p + q = 19 - (1)$$

(ii)
$$S_{xx} = \frac{\sum x^2}{n} - \bar{x}^2 = \frac{61}{5} - 3^2 = 3.2$$
$$S_{xy} = \frac{\sum xy}{n} - \bar{x}\bar{y}$$
$$= \frac{32 + 2p + 6q}{5} - 3\left(\frac{11 + p + q}{5}\right) = \frac{3p - q - 1}{5}$$
$$2.5 = \frac{S_{xy}}{S_{xx}} \rightarrow \frac{3p - q - 1}{5} = 2.5(3.2)$$
$$3p - q = 41 - (2)$$
$$(1) + (2): 4p = 60, p = 15, q = 4.$$

(iii) $S_{xy} = 8$,
$$S_{yy} = \frac{290}{5} - 6^2 = 22$$
$$r = \frac{S_{xy}}{S_x S_y} = \frac{8}{\sqrt{22(3.2)}}$$
$$\approx 0.953 \ (3 \ d.p)$$

(iv)(a) The new regression line is $y = 0.25x - 1.5$.

(b) PMCC is still 0.953. (PMCC is unaffected by coding)
Note: You do not need to show calculation here.

References:

Pure Mathematics

1. CIE Further Mathematics Paper 1 Problems (as quoted)
2. Sketching of Graphs: Wolfram Alpha
3. Calculus: some concepts and proofs are taken from tutorial.math.lamar.edu
4. Elementary Linear Algebra, 10th Edition by Howard Anton
5. Differential Equation Examples: Edexcel FP2 Mathematics Textbook
6. Definitions: from Wolfram MathWorld and Wikipedia.org
7. Oxford Fajar Ace Ahead Series: for coming up with the integration by parts acronym

Mechanics

1. CIE Mechanics 1 by Douglas Quadling (P4 Basic Knowledge)
2. CIE Mechanics 2 by Douglas Quadling (P5 Basic Knowledge)
3. OCR Mechanics 1 by Douglas Quadling (Momentum and Impulse)
4. OCR Mechanics 2 by Douglas Quadling (Momentum and Impulse)
5. OCR Mechanics 3 & 4 by Douglas Quadling (All Other Topics)
6. CIE Mathematics Paper 5 Past Year Problems: As quoted in the notes
7. CIE Further Mathematics Paper 2 Past Year Problems: As quoted in the notes

Statistics

1. Probability and Statistics 1 by Chambers, Crawshaw and Balaam (P6 Basic Knowledge)
2. Probability and Statistics 2 by Chambers, Crawshaw and Balaam (P7 Basic Knowledge)
3. OCR Statistics 3 & 4 by Jane Miller (Continuous Random Variables)
4. A Concise Course in Advanced Level Statistics, Crawshaw and Chambers (All Other Topics)
5. CIE Mathematics Paper 7 Problems: As quoted in the notes
6. CIE Further Mathematics Paper 2 Past Year Problems: As quoted in the notes
7. Definitions: From Wolfram Math World http://mathworld.wolfram.com

Disclaimer: Some of the images and pictures used to write this set of notes are from Google and I do not own copyright for any of them.